ESD-TR-77-273

BEST AVAILABLE COPY

AD A042817

Project Report

S. B. Nuttall

Linear Prediction and Maximum Entropy Spectral Analysis for Radar Applications

24 May 1977

Prepared for the Ballistic Missile Defense Program Office
by the Advanced Sensors Directorate under Electronic Systems Division Contract F19628-76-C-0002 by

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Approved for public release; distribution unlimited.

DDC
DEC 15 1977
C

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. This program is sponsored by the Ballistic Missile Defense Program Office, Department of the Army; it is supported by the Defense Nuclear Agency under Air Force Contract F19628-78-C-0002.

This report may be reproduced to satisfy needs of U.S. Government agencies.

The views and conclusions contained in this document are those of the contractor and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the United States Government.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Raymond L. Loiselle, Lt. Col., USAF
Chief, BMD Lincoln Laboratory Project Office

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

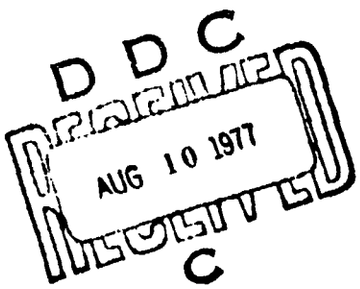
LINEAR PREDICTION AND MAXIMUM ENTROPY SPECTRAL ANALYSIS FOR RADAR APPLICATIONS

S. B. BOWLING

Group 35

PROJECT REPORT RMP-122
(Radar Measurements Program)

24 MAY 1977



Approved for public release; distribution unlimited.

LEXINGTON MASSACHUSETTS

ABSTRACT

For most applications in radar data processing, the Fourier transform performs satisfactorily. However, other methods of spectral analysis can offer some advantages when a data set is too short for a Fourier transform to resolve or detect important spectral features. This report describes one alternative technique, maximum entropy spectral analysis (MESA), and suggests possible radar applications including range-Doppler sizing and the coherent measurement of range rate. Practical examples demonstrate an improvement in velocity resolution and cross-range resolution. Computer codes are listed that calculate MESA power spectra for a real or complex discrete time series.

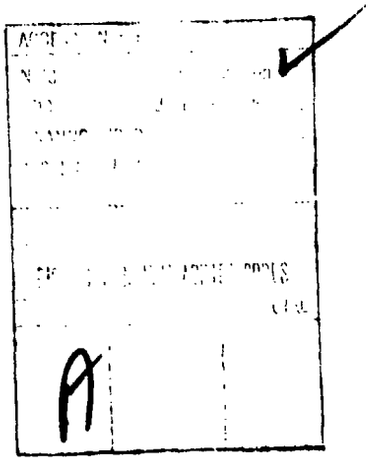


TABLE OF CONTENTS

1. INTRODUCTION

The interpretation of coherent radar signals often involves the calculation of a power spectrum or a periodogram that describes the frequency content of data. The conventional Fourier approach, based on the work of Wiener (1950) and of Blackman and Tukey (1959), relates the autocorrelation function of a signal and its power spectrum through the Fourier transform. Cooley and Tukey (1965) popularized the Fourier approach with the computationally efficient fast Fourier transform (FFT), which has dominated the analysis of radar data until the present time.

Notwithstanding the speed and mathematical elegance of the Fourier transform with its entourage of weighting functions and tapering schemes, it is troubled by several unavoidable limitations that become serious as the temporal length of the data set shortens. However, alternative methods of spectral analysis do exist that have not as yet found wide application in radar. Among these techniques is "maximum entropy spectral analysis" (MESA), an outgrowth of the predictive deconvolution filtering techniques developed by geophysicists for oil exploration. The primary purpose of this report is to introduce MESA to radar as a means of improving the velocity resolution and cross-range resolution currently limited by conventional Fourier concepts.

Section 2 briefly identifies some of the problems with the Fourier approach to spectral analysis. Section 3 describes the MESA procedure for which more detailed derivations are given in the appendices. Comparing conventional and MESA spectra, Section 4 treats several radar applications to include the measurement of range rate and the range-Doppler sizing of hard bodies. Many other applications are feasible. In principle, MESA can replace a Fourier transform wherever the latter may occur, whether a transformation is to be made from the time domain to the frequency domain, or from the frequency domain to the time domain.

2. CONVENTIONAL FOURIER TECHNIQUES

Conventional methods of spectral analysis make unrealistic assumptions about the nature of data outside of the observation interval. The periodogram approach (Jones, 1965) projects the data set as an infinite periodic repetition of itself, while the familiar autocorrelation approach (Blackman and Tukey, 1959) assumes a zero extension. These schemes attempt to lessen the effects of the finite length of a data set upon Fourier transformation.

One may think of a data set truncated in the time domain as the product of an infinite data set and a window function that has non-zero amplitude only within the observation interval. The measured power spectrum of this product is the convolution in the frequency domain of the true spectrum of the data set (which one would like to have measured) and the spectrum of the window function itself. The measured spectrum is unavoidably a distortion of the true spectrum; the convolution allows the calculation of the spectrum at one frequency to be contaminated by energy at all other frequencies (loosely termed "leakage through the sidelobes" of the window function). The measured power spectrum can even take on physically meaningless negative values if negative lobes in the window spectrum are being convolved with strong frequency components in the true spectrum.

For example, a rectangular window (the default window) has a spectrum of the form

$$W(f) = \sin(\pi fT)/\pi fT$$

where T is the length of the data set. The first zero occurs at $f = 1/T$ and, by rule of thumb, frequency components in the true spectrum spaced more closely than 1/T are not easily resolved in the measured spectrum owing to the smearing caused by convolution. The frequency resolution of the measured spectrum is therefore $\delta f \approx 1/T$.

In radar applications, frequency shift and velocity are related by the equation

$$v = \lambda f/2$$

so that velocity resolution becomes the familiar result

$$\delta v = \lambda \delta f/2 = \lambda/2T.$$

Here λ is the radar wavelength. We observe that this limitation on the resolution of velocity is a result of using conventional Fourier spectral analysis and does not necessarily apply if another method of spectral analysis is used.

The rectangular window is sometimes replaced by a tapering function (Hamming, Bartlett, and Taylor windows are a few examples), designed to reduce the effect of the window sidelobes when the window spectrum is convolved with the true spectrum.

The choice of a window always involves a compromise between frequency resolution and the extent to which window sidelobes allow the calculation of the measured spectrum at one frequency to be contaminated by components at all other frequencies.

The Fourier transform loses resolution when the data set is short compared to the periods of the spectral components in the data. Toman (1965) and Jackson (1967) have shown, for example, that most of the energy of a sinusoid will appear near zero frequency in a spectrum calculated with a Fourier transform if the length of the data set is less than 58% of the period of that sinusoid. Interference effects caused by the window sidelobes can also produce spectral shifts as large as 16%.

Basically, the problems from which conventional techniques suffer are caused by the finite length of the data set. While window theory is elegant, it treats only the symptoms of truncation and, in doing so, corrupts perfectly good data with weighting functions. Ables (1974) cites what he calls the "First Principle of Data Reduction":

> The result of any transformation imposed on experimental data shall incorporate and be consistent with all relevant data, and be maximally non-committal with regard to unavailable data.

Conventional Fourier techniques clearly violate this principle: data not in evidence are unrealistically assumed

and true data are changed by weighting functions. Conventional techniques are therefore philosophically objectionable and become less reliable when the data set becomes short – and it is precisely the short data set with which radar is usually concerned.

3. MAXIMUM ENTROPY SPECTRAL ANALYSIS (MESA)

3.1 Deconvolution Filtering

The maximum entropy approach to spectral analysis is a variation of the deconvolution filtering techniques developed by geophysicists for processing seismic signals. A deconvolution filter whitens the spectrum of the signal on which it operates; that is, when convolved with the original signal, it outputs a new signal with a constant (white) spectrum. Mathematically expressed, the convolution of a discrete-time series $x(n)$ with a digital filter with coefficients $a(n)$ is

$$f(n) = \sum_{k=0}^{M} x(n)\, a(n-k). \tag{3-1}$$

The spectrum of $f(n)$, which is the product in frequency space of the spectrum of $x(n)$ and the transfer function (or impulse response function) $A(f)$ of the filter, is a constant:

$$F(f) = \text{constant} = X(f)\, A(f). \tag{3-2}$$

The power spectrum of $x(n)$ is then

$$P(f) = |X(f)|^2 = K/|A(f)|^2 \tag{3-3}$$

where K is a constant. Simply stated, deconvolution filtering involves finding the digital filter that changes the input signal into an output signal with a constant spectrum.

This approach to spectral analysis is also known as the Markov spectrum or the autoregressive spectrum described by statisticians (see, for example, Parzen, 1969). Burg (1967; 1968; 1975) realized that this approach yields the spectrum having the "maximum entropy" (explained below) of all possible spectra that are consistent with the measured autocorrelation function of x(n). Burg also devised methods of efficiently calculating the filter coefficients from which A(f) can be determined.

One advantage of deconvolution filtering is immediately obvious: finding X(f) does not involve a convolution in frequency space with a cumbersome window spectrum that (unavoidably) destroys spectral resolution. The convolution has already taken place in the time domain between the input signal and the digital filter. Therefore, there are no window sidelobes or serious end effects as occur upon conventional Fourier transformation. The truncation of the data set is important only to the extent that enough data must be available to allow the construction of an efficient whitening filter that can reduce the data to a random series.

Other comparisons between the characteristics of Fourier and MESA spectra will be drawn later and are summarized in the appendices. However, a discussion of the philosophy of maximum entropy is first in order.

3.2 The Maximum Entropy Philosophy

Choosing the best spectral-domain representation of a truncated discrete time series, for which only an imperfectly determined autocorrelation function can be calculated, is a major problem in signal analysis. Among the countless spectra that may be consistent with a given autocorrelation function, only one spectrum can be optimal. A set of rules governing that choice must be established.

Jaynes (1957) introduced a method of statistical inference called the "maximum entropy estimate". He showed that information theory (Shannon and Weaver, 1949) provides a criterion for selecting the best statistical description of a process when only a partial knowledge of that process is available. The optimal choice is the one which is maximally non-committal with regard to any missing data, and which is simultaneously constrained to be consistent with all available data. The result is the best estimate that could have been made on the basis of the data at hand. (Actually, Ables's first principle is a restatement of Jaynes's conclusion.)

The term "entropy" is used in an informational sense, such that a measure of the entropy of a process is a measure of our ease in coping with data that we do not have. We must make no assumptions and impose no constraints that cannot be justified directly with the data that we do have.

What is needed is an expression for the entropy of a process in terms of its spectrum. That expression is then maximized subject to the constraints imposed on the spectrum by the available data.

Shannon and Weaver (1949) have shown that the appropriate expression for the entropy of a process in terms of its power spectrum P(f) is

$$E = \int_{-\infty}^{\infty} \log P(f)\, df. \qquad (3\text{-}4)$$

The values of the autocorrelation function calculated from the available data comprise the constraints on P(f). Ideally, every value $\phi(m)$ of the autocorrelation function calculated at lag m is related to P(f) by

$$\phi(m) = \int_{-\infty}^{\infty} P(f) \exp\,[-i2\pi mf\Delta t]\, df \qquad (3\text{-}5)$$

where Δt is the time spacing of the data.

Maximizing Eq. (3-4) subject to the constraints of Eq. (3-5) (one for each known value of $\phi(m)$) becomes a problem in the calculus of variations. In the case where the $\phi(m)$ are equally spaced and centered at zero lag, P(f) may be determined with Lagrange multipliers (Edwards and Fitelson, 1973).

However, a more direct approach, which is equivalent to the formal use of Lagrange multipliers but which allows a more tangible appreciation of the MESA procedure, is the determination of the deconvolution digital filter that transforms a given time series into a random series with a white spectrum. The next section and Appendix I outline this calculation from which A(f) in Eqs. (3-2) and (3-3) is found.

3.3 Prediction-Error Filtering

The whitening of a discrete-time series can be done with a prediction-error filter (Peacock and Treitle, 1969; Makhoul, 1975). A complex measurement x(n) is approximated by the weighted average of the preceding M terms, or by the weighted average of the successive M terms. The former is a <u>forward</u> prediction, and the latter is a <u>backward</u> prediction:

$$x_f(n) = \sum_{k=1}^{M} a_{M,k}\, x(n-k) \qquad (3\text{-}6)$$

$$x_b(n) = \sum_{k=1}^{M} a^*_{M,k}\, x(n+k) \qquad (3\text{-}7)$$

where the asterisk denotes complex conjugation, and $a_{M,k}$ is the kth filter coefficient out of a total of M coefficients. The prediction filter is a linear combination of M data points that

predicts the data point immediately preceding (or following) the M points. The errors in the forward or backward prediction of a given data point x(n) are

$$e_f(n) = x(n) - x_f(n) \tag{3-8}$$

$$e_b(n) = x(n) - x_b(n) \tag{3-9}$$

and, for the linear prediction to be optimal, the errors should be minimized simultaneously in a least squares sense. Therefore, we minimize the total error power

$$P_M = \sum_n \left[e_f(n)^2 + e_b(n)^2 \right] \tag{3-10}$$

with respect to the M coefficients. P_M is the power of the output error series left behind after the x(n) have been predicted. The spectrum of the output series is a constant because, if the spectrum contained any recognizable frequency components, then we could use the knowledge of those components to improve the prediction by editing the filter coefficients until the spectrum does become white.

Therefore, the prediction filter can be used to transform the data x(n) into another series (the error series) that has a constant spectrum: each point x(n) is replaced by the error in its prediction. The prediction-error filter is represented by

Eqs. (3-8) and (3-9) which use the prediction filter. The output of the former is the error in the prediction of a known measurement; the output of the latter is simply the prediction itself.

Van Den Bos (1971) has shown that the prediction-error (P-E) filter is equivalent to the least-squares fitting of a discrete-time all-pole model to the data. Anderson (1974) extended this work in real form, and a complex formulation of Anderson's algorithms, which calculate the P-E filter coefficients, is derived in Appendix I.

The impulse response function of the P-E filter is its z-transform

$$A(f) = 1 + \sum_{k=1}^{M} a_{M,k}\, z^{-k} \qquad (3\text{-}11)$$

$$z = \exp\,[i2\pi f\Delta t]. \qquad (3\text{-}12)$$

There is one coefficient $a_{M,k}$ for each of the M constraints put on the power spectrum during entropy maximization.

The choice of M is somewhat arbitrary, although Akaike (1969a, b; 1970) has argued that the length M of the filter should be chosen so as to minimize the error power P_M. An M should be chosen so that increasing the filter length to M + 1

no longer significantly reduces the error in prediction; that is, P_{M+1} is not much smaller than P_M. Of course, M cannot exceed the number of data points. (See the review article by Ulrych and Bishop, 1975.) Using too small an M results in a highly smoothed spectrum, obviating the improved resolution capabilities of MESA; using too long a filter allows noise to introduce spurious detail into the spectrum. A reasonable initial value of M is one fifth the number of the data points of the input sequence.

3.4 Power Spectrum and Power Spectral Density

Both Lacoss (1971) and Burg (1975) have pointed out that the MESA power spectral density function is a better measure of the amount of power in a small bandwidth than is the power spectrum. The two are related by

$$PSD(f) = \int_{f-\delta f/2}^{f+\delta f/2} P(f)\, df/\delta f \qquad (3\text{-}13)$$

where PSD(f) is the power spectral density, P(f) is the power spectrum (Eq. (3-3)), and δf is the small bandwidth over which P(f) is integrated. Since P(f) is calculable at any frequency, A(f) being an analytic and continuous complex polynomial, Eq. (3-13) is easily implemented using numerical integration (see appendices).

3.5 MESA and Non-Linearity

One attractive property of conventional Fourier techniques is linearity: the power spectral density of the sum of several signals is the sum of the power spectral densities of the individual signals. Moreover, the square root of the Fourier PSD(f) gives a reliable amplitude spectrum of the input signal.

Unfortunately, MESA is ultimately not a linear technique. Indeed, Eq. (3-3) is the inverse of the square of a complex polynomial. Therefore, a comparison of relative power at different frequencies can be misleading. Parseval's theorem notwithstanding, the integral of the MESA PSD(f) may not equal the total power of the input signal.

The value of MESA lies in frequency detection and frequency resolution, as discussed later in Section 3.7. It is not as useful for determining the relative strengths of different frequency components. MESA can be used in conjunction with Fourier techniques, for example, by identifying the important frequency components with MESA and then performing a DFT (discrete Fourier transform) only at those frequencies to estimate their strengths and phases. Section 3.6 offers another alternative in which a Fourier transform can be taken for a data set extended beyond the observation period by the linear prediction filter.

It is worthwhile at this point to mention the non-linear effects of noise on the MESA spectrum. We can rewrite Eq. (3-3) as

$$P(f) = K/|z^{-1}-z_1|^2 \, |z^{-1}-z_2|^2 \ldots |z^{-1}-z_M|^2 \quad (3\text{-}14)$$

$$z = \exp(i\Theta) = \exp(i2\pi f\Delta t)$$

where we have factored A(f) in the denominator. If we plot the poles z_ℓ (ℓ=1 to M) of P(f) in the complex plane, as demonstrated in Fig. 3-1 for M=5, they will all lie within the unit circle. (This is because the set of prediction filter coefficients is "minimum phase", as explained, for example, in Robinson (1967) or Claerbout (1976).) As A(f) is evaluated on the unit circle, the angle $\Theta=2\pi f\Delta t$ varies from $-\pi$ to π. Therefore, f varies between $-1/2\Delta t$ and $1/2\Delta t$, that is, between minus and plus the Nyquist frequency. For a given frequency at point F, each factor in the denominator of Eq. (3-14) is the square of the distance between point F and one of the poles. As f changes and point F passes by a pole near the unit circle, P(f) will exhibit a local maximum.

The frequency components in the input data set will correspond to those poles of P(f) closest to the unit circle. Clearly, there can be no more maxima detected in P(f) than there are poles of P(f) or coefficients of the filter. In the

case where there are actually fewer frequency components in the data than there are poles available for placement, the extra poles can be eliminated by reducing the number of filter coefficients, as M is the degree of the complex polynomial A(f) that is factored to locate the M poles.

Noise can affect the location of the poles in the complex plane. Since Eq. (3-14) involves the inverse squares of the distances between a point on the unit circle and the poles, a small change in the position of a pole that is already near the unit circle can cause a large change in the magnitude of P(f) near that pole. It is for this reason that the amplitude of P(f) may not accurately reflect the true power spectrum. Also, since the PSD(f) is the line integral of P(f) over a small arc on the unit circle (a small frequency interval δf), the amplitude of the power spectral density similarly may not be accurate.

Burg (1975) has recognized that both the peak value and width of an apparent spectral line in P(f) strongly depend on the background noise. However, he maintains that the product of the peak value and line width, which is proportional to the total power in the spectral line, will be estimated accurately if the sampling of the spectrum is fine enough to trace out the shape of the spectrum. In general, the unequivocal detection of a spectral component is much more reliable than an estimate

of its strength in a MESA spectrum. Therefore, locating the poles of P(f) may be sufficient for frequency detection.

Akaike (1969), Baggeroer (1976), and others cited therein have investigated some confidence intervals and the statistical variability of the MESA technique.

The MESA user must decide if the non-linearity poses a problem for his specific application. If the detection of frequency components in a data set is the primary goal, MESA will indeed be useful. However, the procedure may not be totally adequate if precise amplitude and phase information is also required. In that case, a DFT can be computed at the frequencies already identified by the MESA procedure.

3.6 Linear Prediction and the Fourier Transform

The MESA procedure is linear up to the point at which the prediction filter is calculated. The filter itself is a linear operator that, as we have seen, uses a weighted average of M data points to estimate the adjacent data points.

We propose the following procedure to take advantage of the linearity of both the prediction filter and the Fourier transform: A prediction filter is calculated from a data set of N_1 points. The filter is then used to extend (or predict) the original data set to a total of N_2 points, outside of the

observation interval in both forward and backward directions. Then a conventional Fourier spectrum is calculated with the larger data set of N_2 points.

The data points that are predicted by the filter will have essentially the same spectral composition as the original data, as a knowledge of the spectrum of the original data is contained in the filter coefficients.

Since additional signal is being created and subsequently transformed, the normalization of the power spectrum of the extended data set should be done with respect to the total number of samples N_2. Any weighting window should be applied across the N_2 samples, after the extension has been done.

The deterministic spectral components having been reinforced in the process of prediction, the apparent signal-to-noise ratio is improved; the contribution of noise cannot be predicted and therefore does not appear in the extension. Of course, the quality of the prediction will depend on the extricable determinism and the SNR of the original data. If the original data are pure noise, the poles of P(f) will be near zero (clustered at the origin of the complex plane) and the magnitudes of the predicted points will be very small or zero. Otherwise, the sinusoidal components of the original data dominate the extension.

Linear prediction is less likely to improve the regions of a power spectrum for which a band or continuum of frequencies contains energy. In this case, discrete poles will position themselves in the complex plane in the vicinity of the arc on the unit circle corresponding to that band of frequencies. Some frequencies in the band may be enhanced slightly more than others because the poles are points, each of which is closest to a single point (frequency) on the unit circle. The uneven effect will be small, however, and the poles will move farther away from the arc on the unit circle as the band of frequencies broadens. Indeed, if there is equal energy at all frequencies (noise), the poles cluster near the origin, which is the only point equidistant from every point on the unit circle.

Predicting the data beyond the observation interval is more palatable than assuming such data are zeroes. The prediction is as self-consistent with the original data as is possible with the information at hand. Conventional zeroes would be no more consistent with the original data as would be any other constant randomly chosen.

The appendices contain a FORTRAN program "LNPRED" that linearly extends a complex data set from N_1 to N_2 points. A user-supplied Fourier transform can then be performed on the N_2 points output from this prediction subroutine.

3.7 Improved Frequency Resolution with MESA

Because a data set could, in principle, be extended to infinity, spectra based on linear prediction have a better chance both of separating closely spaced frequency components and of locating them more precisely in frequency space. In this section, we demonstrate that MESA can more accurately estimate the frequency of a single noisy sinusoid, as well as detect two closely spaced sinusoids, than can conventional techniques.

First consider a noisy sinusoid sampled in time:

$$x(n\Delta t) = A \sin(2\pi f_o n\Delta t + \phi) + N(n\Delta t) \qquad (3\text{-}15)$$

where f_o is the basic frequency, ϕ is a phase constant, Δt is the sample spacing, A is the sinusoid amplitude, n is a sample counter, and N is Gaussian noise. We want to calculate the error in measuring f_o, as a function of a given signal-to-noise power ratio $(A/N)^2$ and of a fixed number of data samples, when either MESA or conventional techniques are used.

We assemble an ensemble of 100 data sets $\{x(n\Delta t)\}$ by making 100 random draws on ϕ between 0 and 2π. Each data set has the same number of samples and the same signal and noise amplitudes. A spectral transform is done on each of the 100 data sets in an attempt to extract f_o. Then the standard deviation of the 100-member set $\{f_{measured} - f_o\}$ is calculated,

and σ/f_o becomes a measure of the precision with which we can find f_o for a given SNR and number of samples. We can construct a family of curves by holding the SNR constant and varying the number of samples, or vice versa.

As long as $f_o < 0.5/\Delta t$, the sinusoid is properly sampled and the actual values of f_o and Δt are irrelevant.

Figures 3-2(a-b) compare the expected error in measuring f_o for a data set that is processed with either MESA or an FFT. We have taken $f_o = 0.25/\Delta t$ (four samples per period).

As expected, either an improvement in SNR or an increase in the number of samples reduces the error. However, for a given SNR and length of data set, the error in the MESA estimate of f_o is consistently smaller than the error in the FFT estimate of f_o.

This reduction in expected error demonstrates that conventional Fourier techniques do lose some information owing to the problems discussed in Section 2. Although we do not expect the Fourier transform of a linearly predicted data set to render an error as small as MESA, it should show improvement over the conventional Fourier transform that uses an extension of zeroes.

As an example of the ability of MESA and Fourier techniques to resolve closely spaced frequencies, consider a two-component signal sampled in time

$$x(n\Delta t) = \sum_{m=1}^{2} A_m[\cos(2\pi f_m n\Delta t) + i \sin(2\pi f_m n\Delta t)] + N(n\Delta t) \quad (3\text{-}16)$$

The shaded part of Fig. 3-3 displays 25 samples of the real part of Eq. (3-16), for which $A_1=A_2=$ 2 units, Δt = 0.01 sec, f_1 = −15.3 Hz, and f_2 = −13.3 Hz. Enough white noise is added to give a SNR of 10 dB.

Since only 0.25 sec of data is transformed, conventional Fourier techniques (zero extension) will be unable to resolve frequencies spaced more closely than 1/0.25 = 4 Hz. In Fig. 3-4 the conventional power spectral density shows only one peak, as expected. The MESA PSD of Fig. 3-5, however, detects the presence of both frequencies in the 25 samples; the unequal power amplitudes are the result of non-linearity.

If we extend the 25 samples to 100 samples with a 5-point prediction filter (the extension is the unshaded region of Fig. 3-3), then the conventional Fourier PSD does resolve both frequencies, as is shown in Fig. 3-6. The power amplitudes in Fig. 3-6 are almost equal and yield a more reliable estimate of the relative strengths of the two frequency components.

Thus, we find that both the MESA spectrum of the original data set and the Fourier spectrum of the linearly-predicted data set can resolve closely spaced frequencies with more success than can the traditional Fourier spectrum.

3.8 Additional Use of the z-Transform

The calculation of a power spectrum or a power spectral density is really unnecessary if only frequency information is required. The z-transform A(f) of the prediction filter can be factored to give

$$A(f) = (z^{-1}-z_1)(z^{-1}-z_2) \ldots (z^{-1}-z_M) \qquad (3\text{-}17)$$

as in Eq. (3-14), where z_ℓ (ℓ = 1 to M) are the zeroes of A(f) and $z = \exp(i\Theta) = \exp(i2\pi f\Delta t)$. The magnitudes of the zeroes $|z_\ell|$ are always between zero and unity, and the zeroes with the largest magnitudes are likely to correspond to the frequency components in the data. If we choose some minimum magnitude which z_ℓ must have to be considered as a candidate frequency component ($|z_\ell| > 0.8$, for example), then we can calculate the corresponding frequency since

$$\Theta = 2\pi f\Delta t = \arctan[\mathrm{Im}(z_\ell)/\mathrm{Re}(z_\ell)]. \qquad (3\text{-}18)$$

For radar applications, range rate $\dot{R} = \lambda f/2$ so that

$$\dot{R} = V_{amb} \arctan[\mathrm{Im}(z_\ell)/\mathrm{Re}(z_\ell)]/2\pi \qquad (3\text{-}19)$$

where $V_{amb} = \lambda/2\Delta t$ is the ambiguous velocity. Equation (3-19) is easily interpreted geometrically because one trip around the unit circle is one foldover in velocity. A target's range rate can be estimated accurately in this way without the use of a Fourier transform.

3.9 Additional Comments

This section has presented the essence of the linear prediction – maximum entropy approach to spectral analysis, avoiding many of the mathematical complexities that are treated extensively in the literature and briefly here in the appendices. The reader interested in more detail is referred to the bibliography in the review article by Ulrych and Bishop (1975) and to Burg's (1975) dissertation which explores many topics not discussed here.

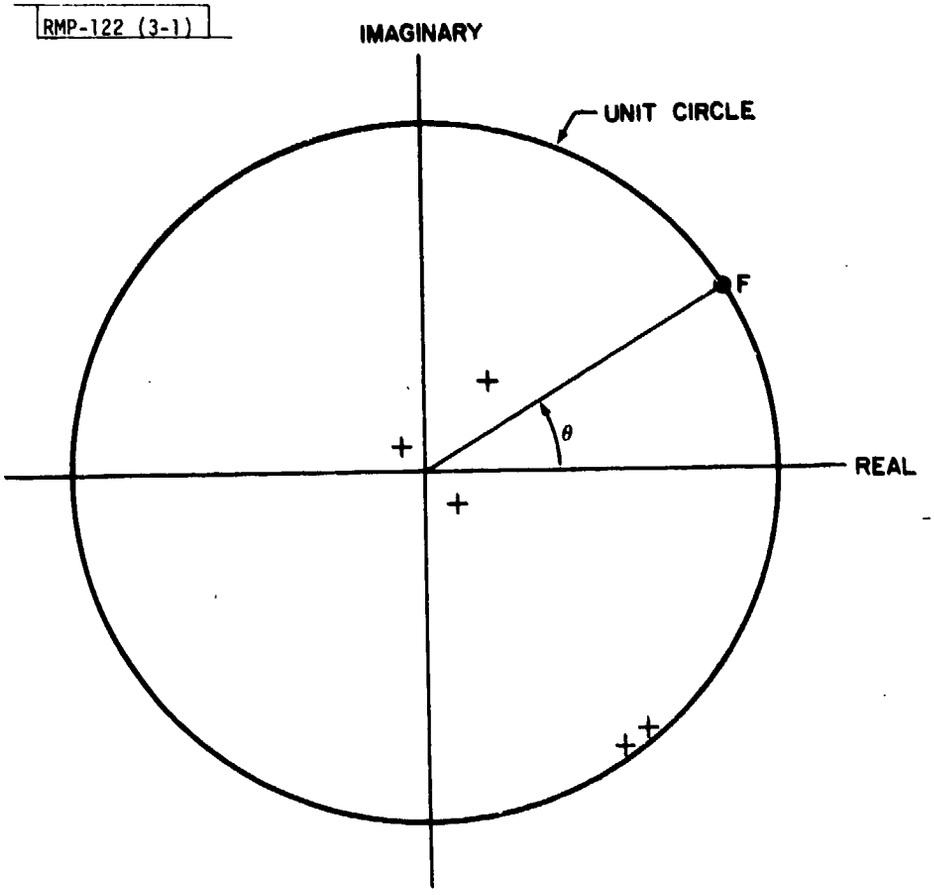


Fig.3-1. Zeroes of the z-transform of a five-point prediction error filter are plotted in the complex plane.

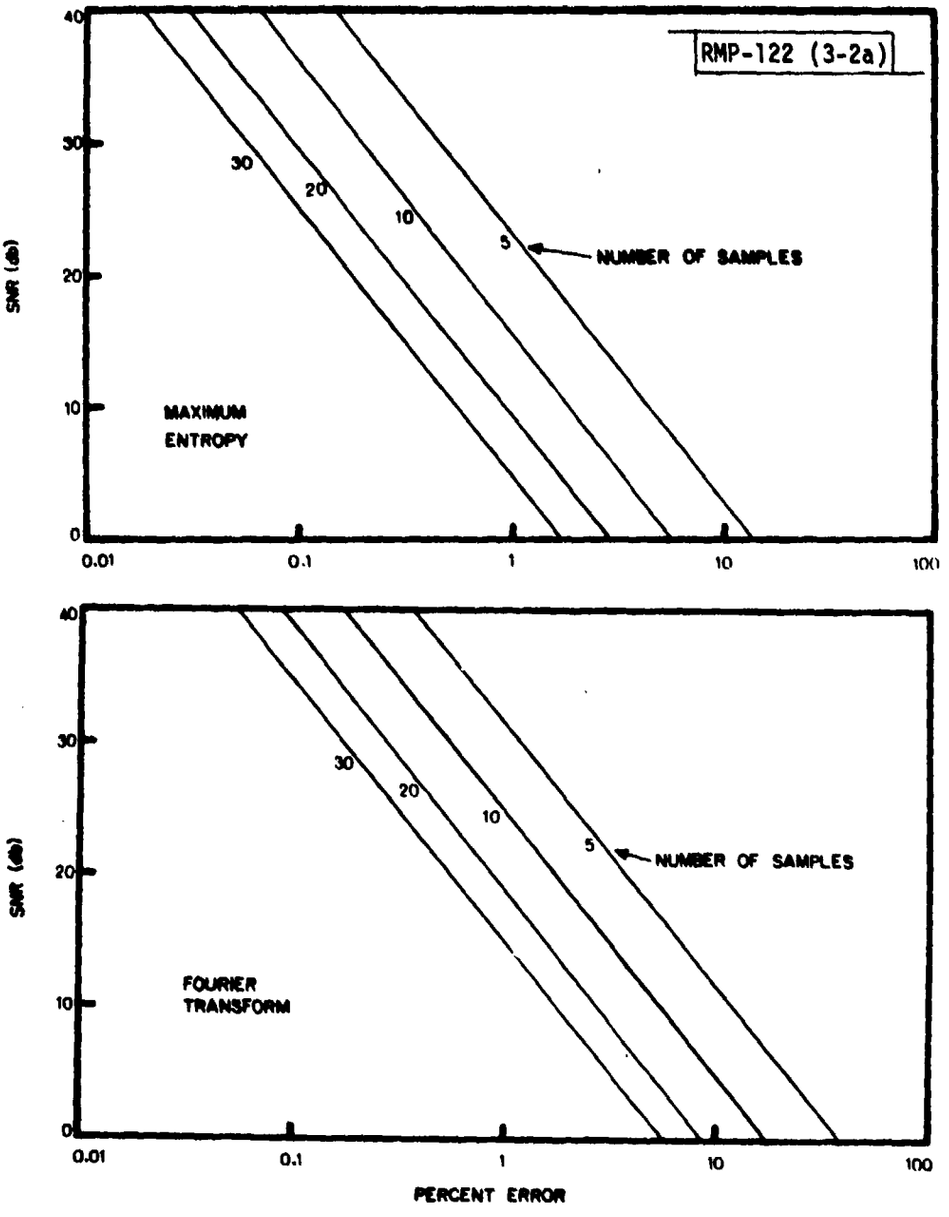


Fig.3-2a. Percent error in estimation of the frequency of a noisy sinusoid as a function of SNR for a given number of data samples.

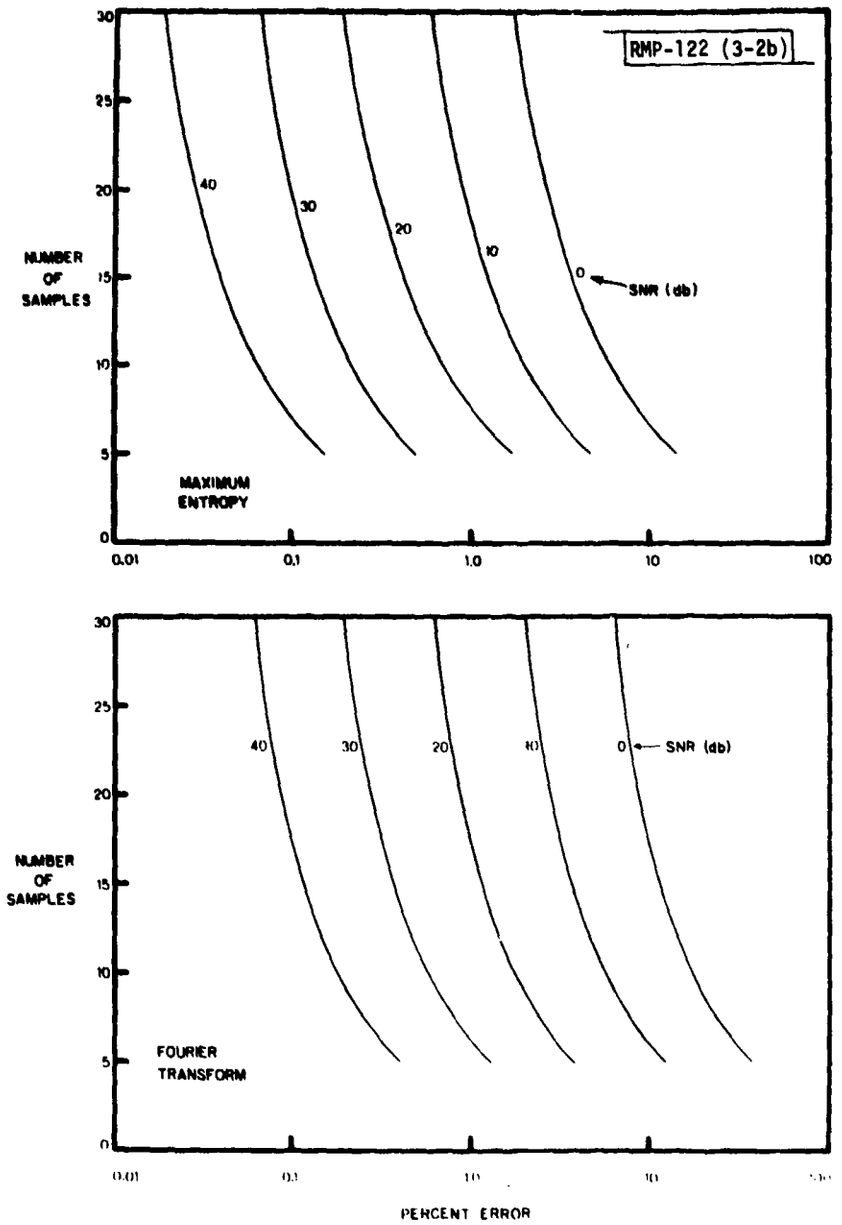


Fig.3-2b. Percent error in the estimation of the frequency of a noisy sinusoid as a function of the number of samples at a given SNR.

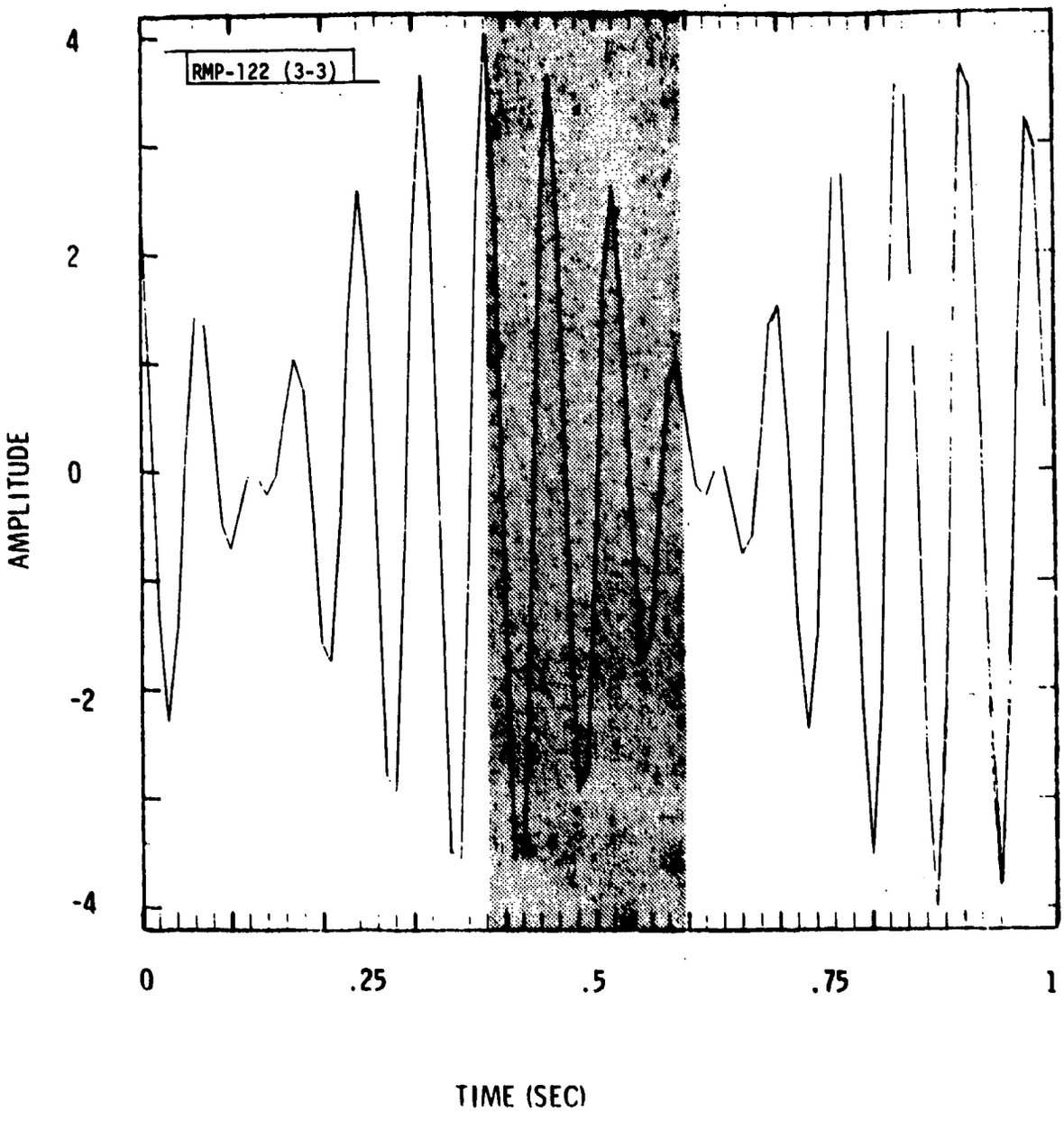


Fig.3-3. A two-frequency signal (−15.3 Hz and −13.3 Hz) is sampled 25 times, every 0.01 second (shaded region), and linearly predicted to a total of 100 samples.

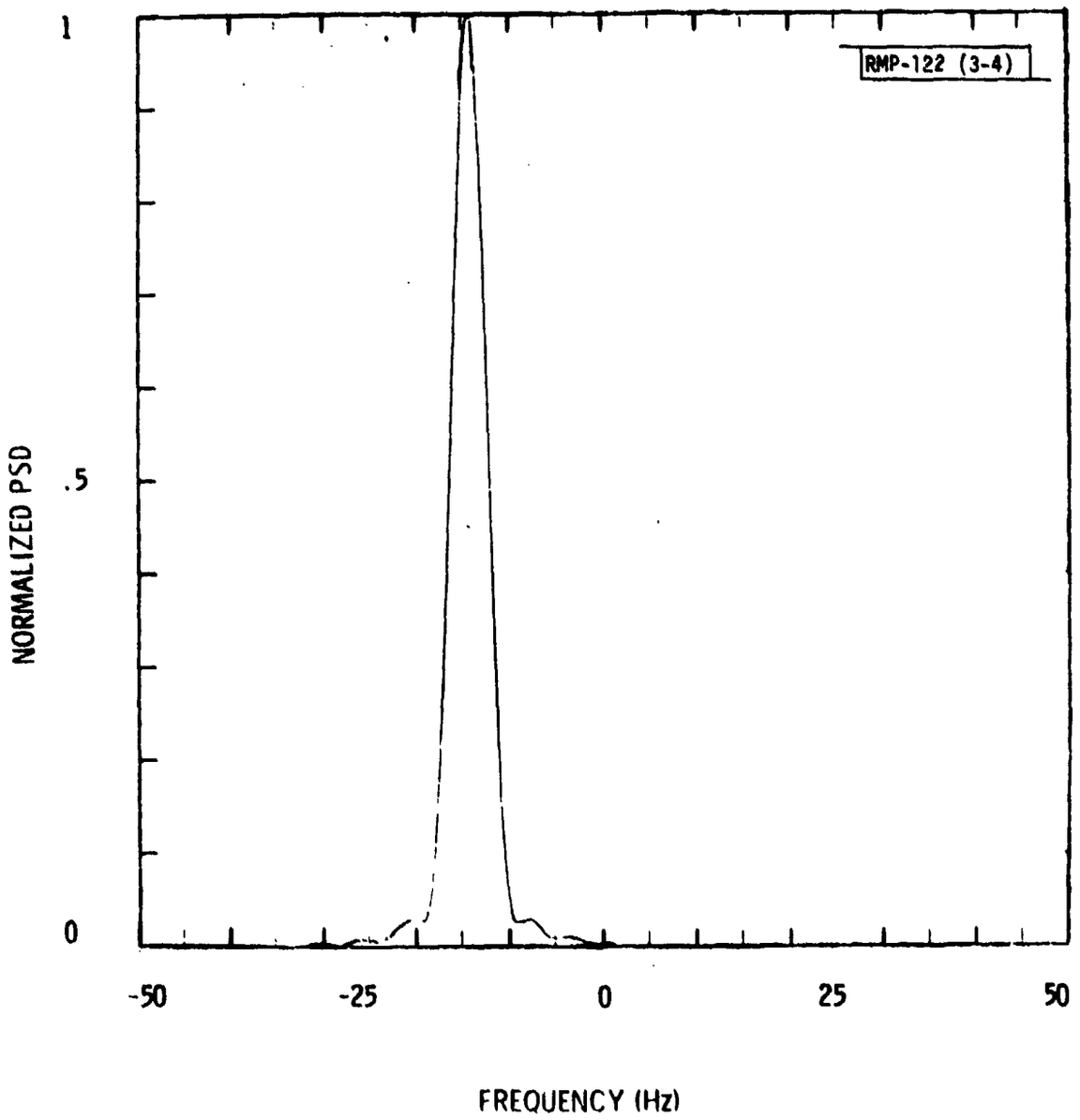


Fig.3-4. Conventional Fourier PSD of shaded region in Fig.3-3.

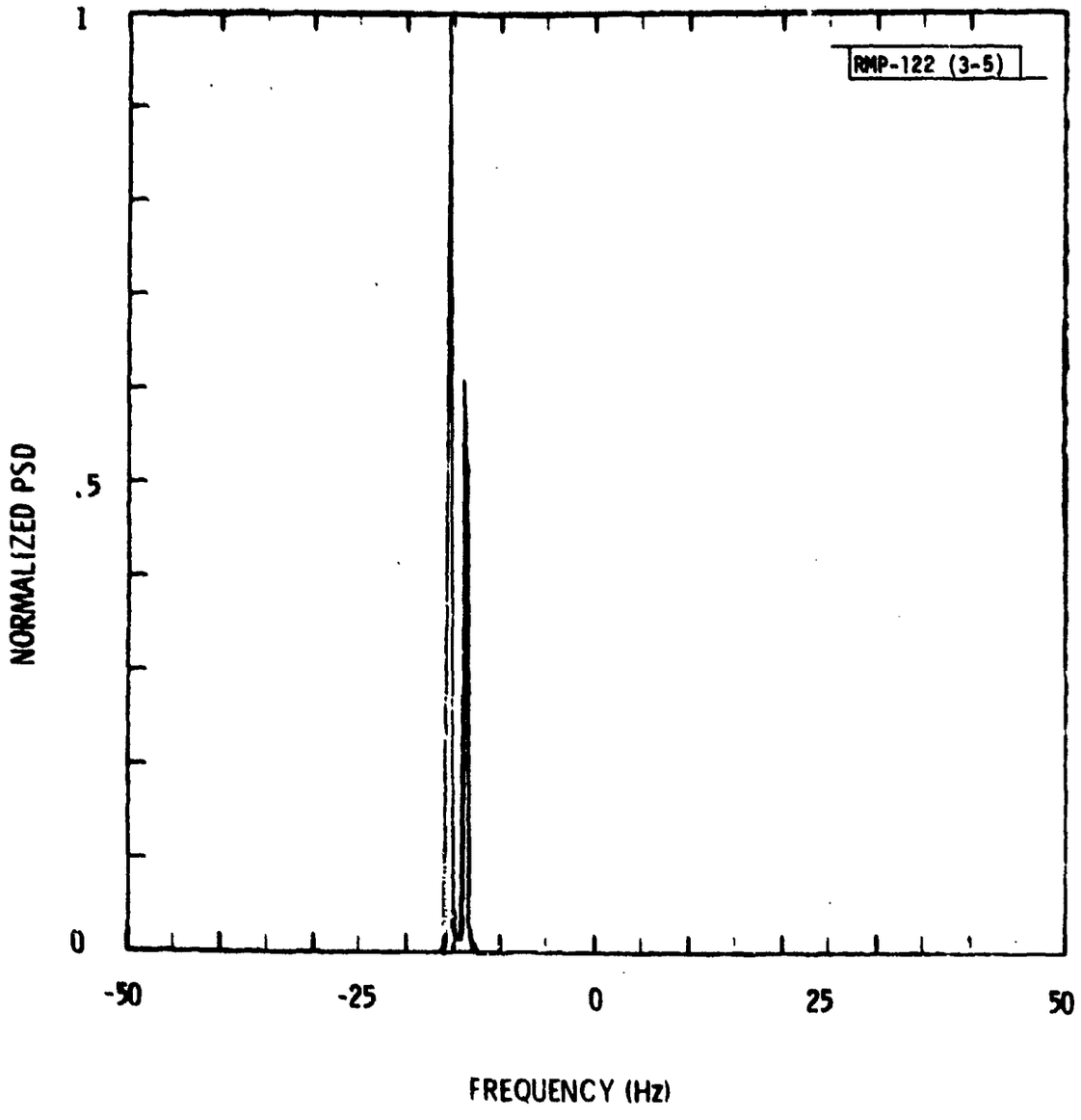


Fig.3-5. MESA PSD of shaded region in Fig.3-3.

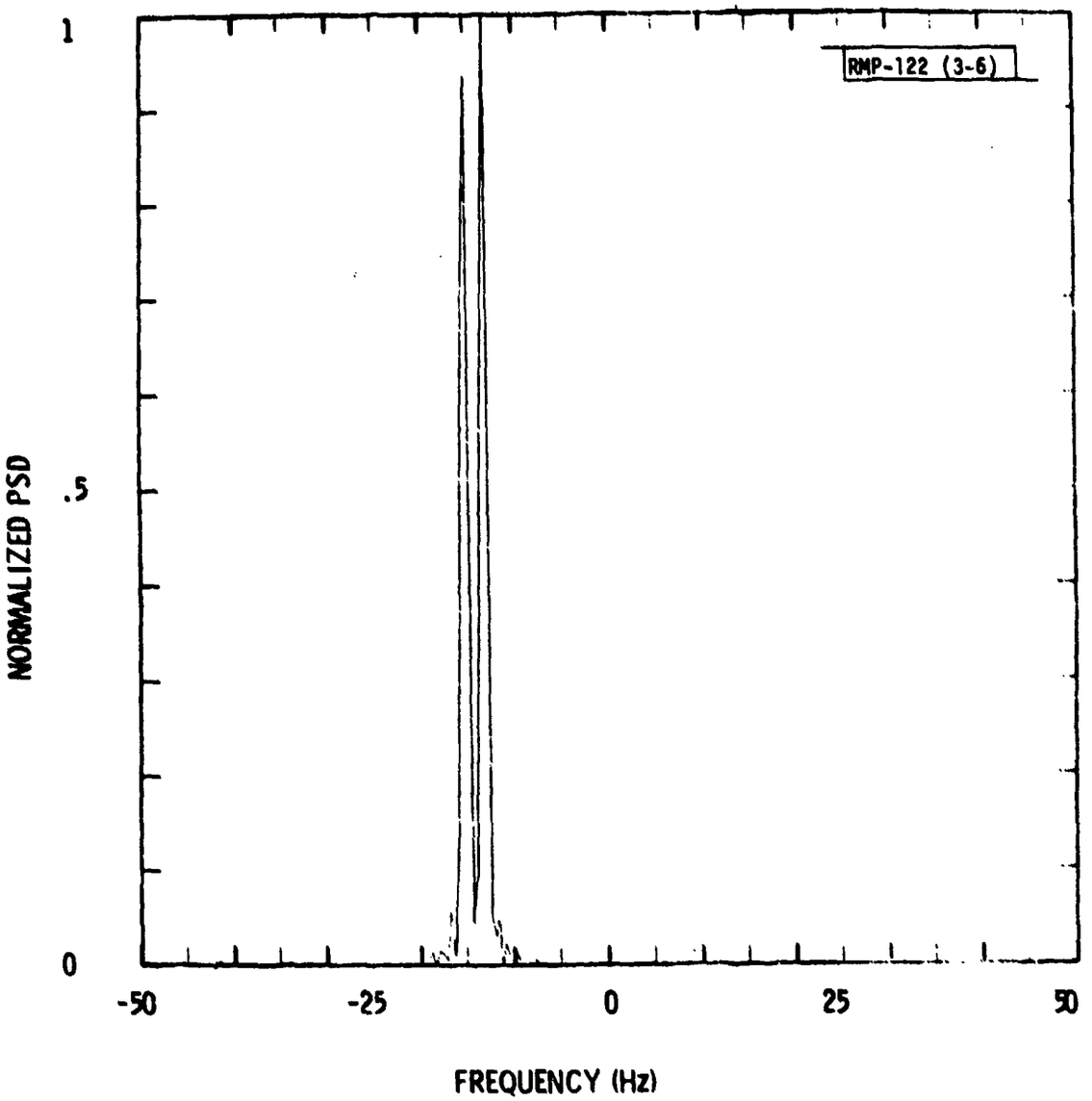


Fig.3-6. Fourier PSD of linearly extended data (all 100 samples) in Fig.3-3.

4. MESA APPLICATIONS FOR RADAR

4.1 Radar Data Preparation

A coherent radar records the amplitude and phase of the energy coming from each range cell along the radar line of sight (RLOS). For a given range cell or range gate, we form a complex sample at time t

$$x(t) = A(t)\cos\Theta(t) + iA(t)\sin\Theta(t) \qquad (4\text{-}1)$$

where A(t) is the radar cross section in volts (or in meters or an equivalent linear unit) and $\Theta(t)$ is the phase. Since the data are recorded pulse by pulse, x(t) is a discrete-time series.

If the object is moving, then $\Theta(t)$ changes in time since phase is governed by

$$\Theta(t) = 4\pi R(t)/\lambda, \qquad (4\text{-}2)$$

where R(t) is the one-way distance between the object and the radar, and λ is the radar wavelength. Pulses must occur often enough that $\Theta(t)$ does not change more than 2π between pulses; otherwise, the phase is aliased and becomes ambiguous.

The amplitude A(t) may change in time if the orientation of the radiation pattern of the object changes with respect to the RLOS, owing to the overall body velocity or to body motion about the center of mass (spin, tumble, precession, etc.).

Hence, $x(t)$ contains information about the scintillation of reflected energy and about the velocities of objects in the range cell. Indeed, $A(t)$ and $\Theta(t)$ represent the resultant vector sum of the individual returns from all unresolved scatterers in the same range cell, each of which contributes a frequency component to $A(t)$, to $\Theta(t)$, or to both.

A non-coherent radar does not record phase, for which case we take $\Theta(t) = 0$ in Eq. (4-1) and $x(t)$ reduces to the real series

$$x(t) = A(t) + i0 = A(t). \qquad (4\text{-}3)$$

In this section, we present examples of the use of Fourier and MESA algorithms to measure body velocity (range rate) and to perform range-Doppler sizing of an object with motion about its center of mass.

4.2 Measurement of Range Rate

The projection of the velocity vector of an object along the RLOS is its range rate, which can be estimated from a coherent data set by the spectral analysis of $x(t)$. For the moment we assume $A(t)$ is a constant and temporal changes in $x(t)$ can be ascribed to changes in $R(t)$ [Eqs. (4-1) and (4-2)].

The frequency axis of the PSD function can be scaled in velocity since $v = \lambda f/2$. One ambiguous velocity interval spans

the Nyquist range from $-1/2\Delta t$ to $1/2\Delta t$, such that $V_{amb} = \lambda(2/2\Delta t)/2 = \lambda/2\Delta t$. An object may be seen to fold over as the range rate, which is measured modulus V_{amb}, moves from one ambiguous interval to another.

This effect is shown in Fig. 4-1 where the range rate of an object observed by a UHF radar gradually changes in time ($\lambda = 0.69$ m, $\Delta t = 1/160$ sec). Here, consecutive data sets of 16 pulses are transformed with a conventional FFT. A zero extension to 256 samples begins at sample 17 and a rectangular window is used. There is a 50% overlap of data from line to line, 8 old pulses dropped and 8 new pulses added each time a new PSD is calculated and plotted.

The main lobe of each power spectral density function locates the range rate of the object during the processing interval T. Here T = 0.1 second. The third dimension, a linear scale from 0 to 1, shows the sidelobe structure of the $\sin(\pi fT)/(\pi fT)$ window spectrum that has been convolved with the true spectrum (a delta function at the true range rate).

We can extend each set of 16 pulses to 48 pulses with a 4-point linear prediction filter before Fourier transformation, as if we were processing 0.3 second of data instead of 0.1 second. Figure 4-2 shows a marked reduction in sidelobe levels and a proportional narrowing of the main lobe. A zero extension

to 256 samples and a rectangular window are used, although the zero extension here begins at sample 49. A 50% overlap of original data points allows Figs. 4-1 and 4-2 to be compared line-by-line.

Figure 4-3 displays the analogous MESA results. The sharpness of the peaks and the absence of sidelobes are two immediately apparent features. As we have seen earlier, the error in locating the position of each peak is considerably less with MESA than with the conventional techniques. Rather than computing the full MESA PSD, the "largest zero" procedure suggested in Section 3.8 could have been used to locate the peaks in Fig. 4-3 with the same precision.

Even when the data set is short and only a few radar pulses are processed, conventional Fourier techniques may perform satisfactorily if there is only one velocity to be estimated. However, when the data are limited, the conventional techniques are less able to detect closely-spaced velocities of multiple objects unresolved in range. For this application the linear prediction and MESA algorithms may prove useful.

4.3 Range-Doppler Sizing

The size estimation and imaging of a hard body rely on cross-range (Doppler) measurements made in each of the range

cells that overlay the target. For example, if an object is spinning, then each frequency component f_i (that is, each velocity component v_i) in the radar return is proportional to the distance r_i from the spin axis at which a scattering center on the target is located. That is,

$$v_i = \lambda f_i/2 = \omega_{spin} r_i \sin\Omega \qquad (4\text{-}4)$$

where λ is the radar wavelength, ω_{spin} is the spin angular frequency, and Ω is the aspect angle between the spin axis and the RLOS. We seek the frequency components f_i. With a knowledge of ω_{spin} and Ω, a range-Doppler image can be constructed from a collage of the cross-range plots from all range cells containing the object.

Figure 4-4 displays the evolution of the conventional Fourier PSD for the radar return coming from the base of a spinning and precessing conical target. The wavelength λ is 5.3 cm, Δt is 0.01 sec, and 32 pulses are processed at once with a 50% overlap of pulses from line to line. Because the spin period is 0.5 second and 0.32 second of data is being transformed per line, the velocity spectrum is not time stationary and, therefore, becomes smeared within a band of frequencies thrcugh which many scattering centers move during the processing interval. The band is modulated by the precessional motion, which changes the aspect angle in time. However, the base

radius can be estimated with Eq. (4-4) and a knowledge of the spin rate and time-dependent aspect angle. The edges of the precessional envelope correspond to the horizons of the base that are spinning toward and away from the radar, so a lower limit on the base radius can be made with Eq. (4-4) where the velocity spectrum has its greatest width.

Analogous to Fig. 4-4 is Fig. 4-5, for which the 32 pulses are extended to 96 before Fourier transformation. Because the data are not initially time-stationary owing to the long processing interval, there is not a significant improvement. Use of the MESA procedures in Fig. 4-6, however, does suppress the sidelobes and allows a clearer definition of the precessional envelope.

A reduction in the length of the data set makes time-stationarity approachable. When the processing interval is small, the scattering centers on the target base cannot exhibit as large a change of range rate and hence severely smear the velocity spectrum. However, conventional Fourier techniques have poorer resolution as the data set becomes shorter; in the effort to detect single frequency components in approximately time-stationary data, we paradoxically lose the ability to resolve them.

For example, Fig. 4-7 shows conventional Fourier spectra when only 8 pulses are processed at a time, as opposed to the 32 pulses used earlier (a 50% overlap is retained). Even though Fig. 4-7 is centered where the precessional envelope necks down (about t = 7.5 seconds in Fig. 4-4), there is no distinct indication that the target is precessing.

If, however, the 8 pulses are linearly extended with a 3-point filter to 24 pulses before Fourier transformation, the precessional envelope becomes evident in Fig. 4-8. In this example, linear prediction has provided the same information about the dynamic motion of the body with one-fourth the amount of data previously used.

Analogous MESA spectra are shown in Fig. 4-9. The absence of sidelobes more clearly reveals the precessional envelope. Moreover, there are suggestions of the paths of individual scattering centers where the spectral peaks sweep diagonally from right to left as time increases (see arrows). An ability to track individual scattering centers would be sufficient to estimate, for example, the spin rate of a target unambiguously.

On the one hand, any of the three methods of Doppler processing seems to provide essentially the same spectral information if the data set is long or if it is not time-stationary. The spectra computed with the linear prediction

and MESA algorithms are, however, less cluttered with sidelobes to allow better definition of the qualitative features in the spectra.

On the other hand, the linear prediction and MESA algorithms can analyze short data sets with more success than can conventional Fourier techniques. Features which the latter may fail to resolve may be detected by the alternative techniques.

4.4 Other Radar Applications

Only two applications have been mentioned here that, nevertheless, demonstrate the utility of predictive deconvolution concepts. Other applications might include

(1) Radar metrics – updating Kalman filters with fewer radar pulses;

(2) Drag measurements – measuring target range rates in less time and with more precision;

(3) Discrimination in clutter – detecting objects in velocity space that scatter weakly compared to the background noise;

(4) Pulse compression – improving range resolution when frequency data are digitally transformed into the time or range pulse shape.

This list is by no means exhaustive. The potential user must decide if the linear prediction and MESA procedures afford an advantage for his own application. In any case, these new techniques are well worth trying and offer an alternative to the conventional Fourier transform.

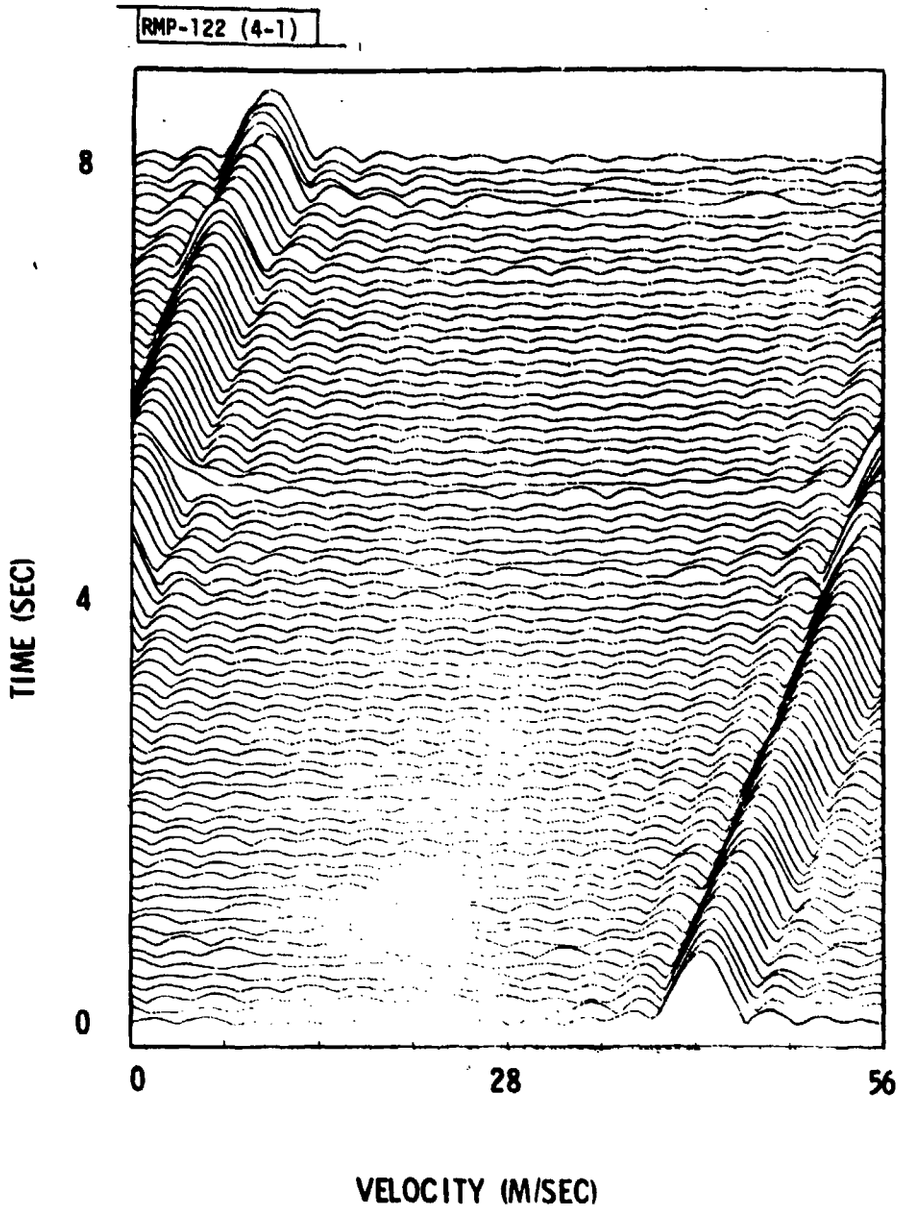


Fig.4-1. Range rate history calculated with conventional Fourier transform.

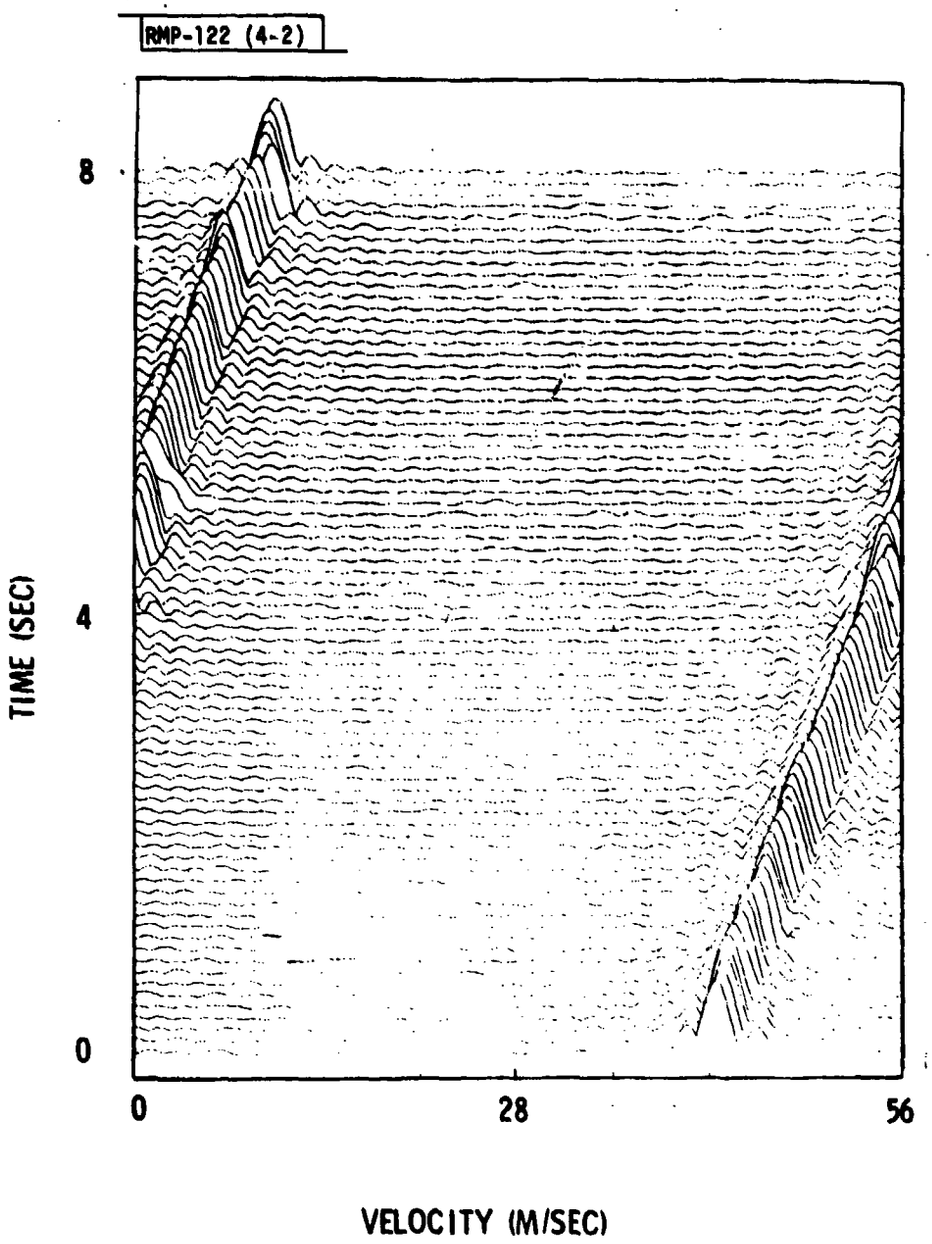


Fig.4-2. Range rate history calculated with a Fourier transform after data have been linearly extended.

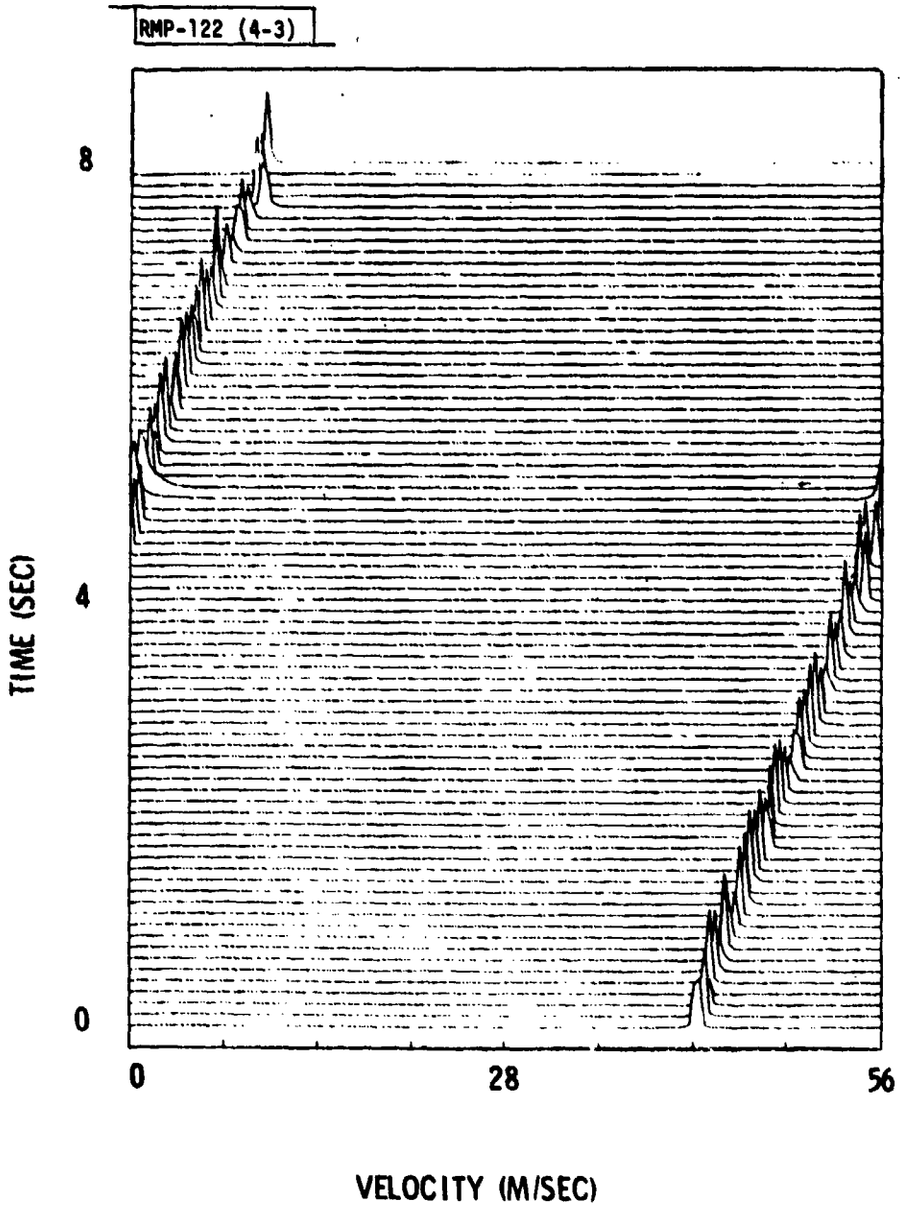


Fig.4-3. [illegible]ge rate history calculated with MESA.

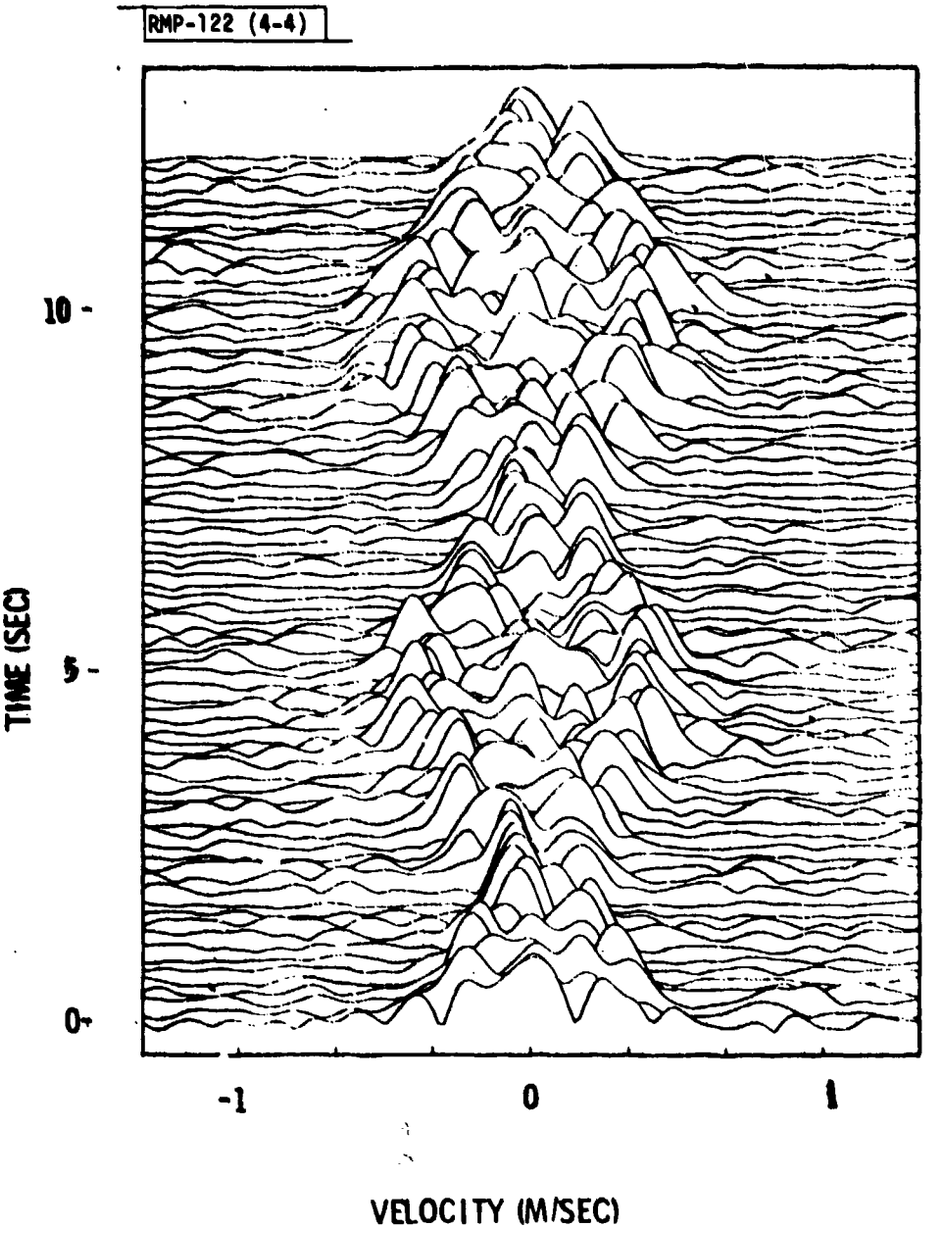


Fig.4-4. Doppler history of base range cell of spinning and precessing cone using conventional Fourier transformation of 32 samples.

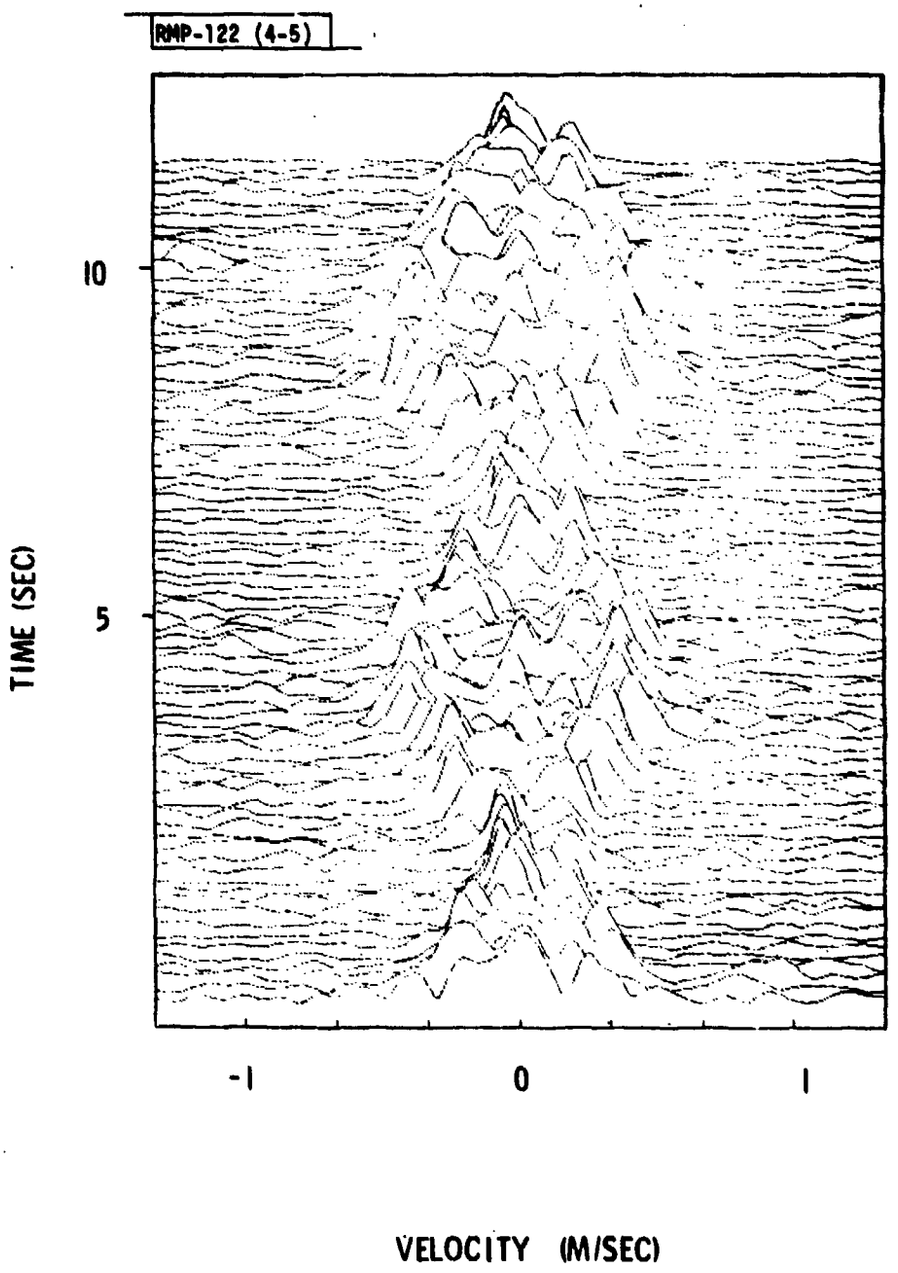


Fig.4-5. Similar to Fig.4-4, but data are linearly extended before Fourier transformation.

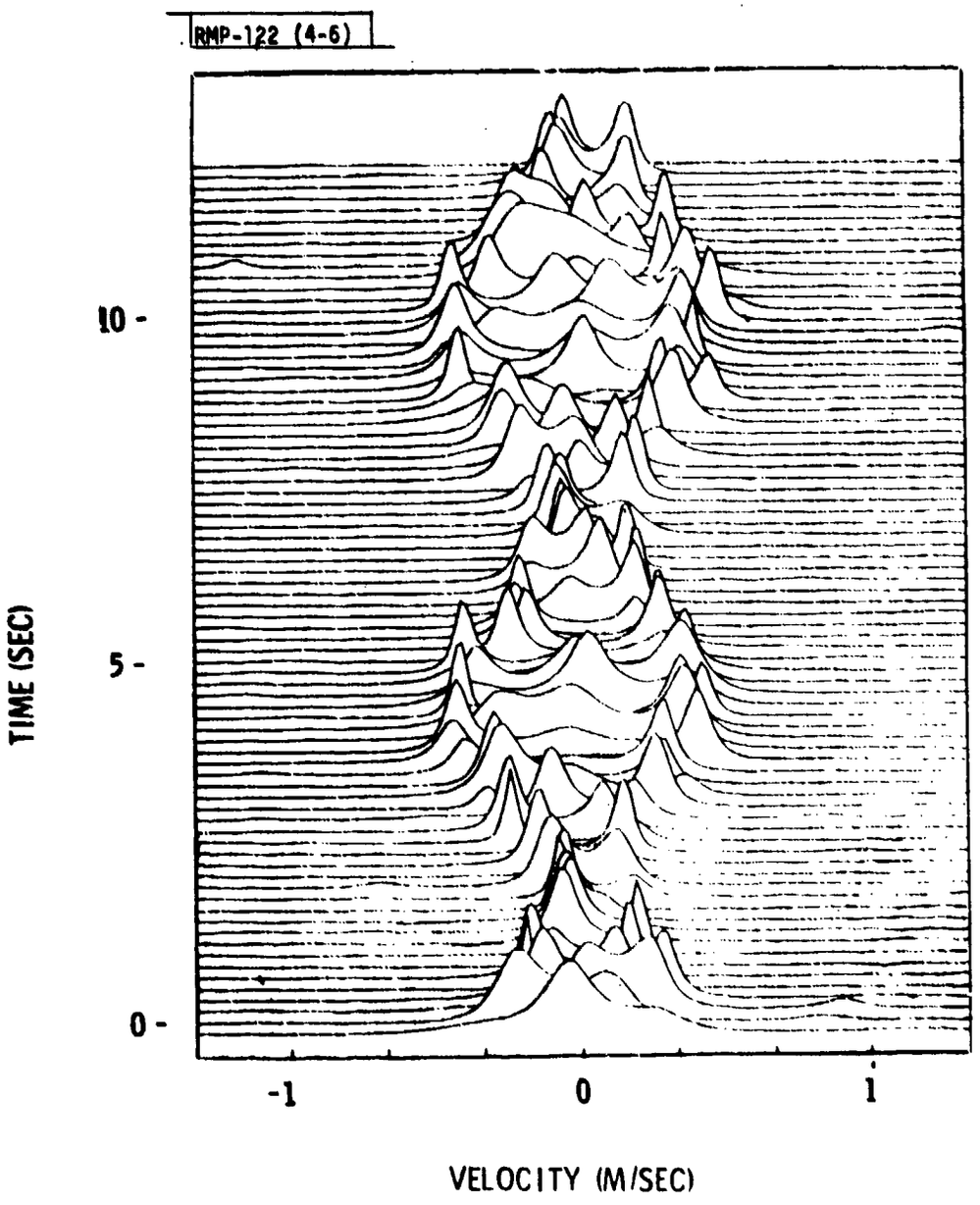


Fig.4-6. MESA Doppler history analogous to Fig.4-4.

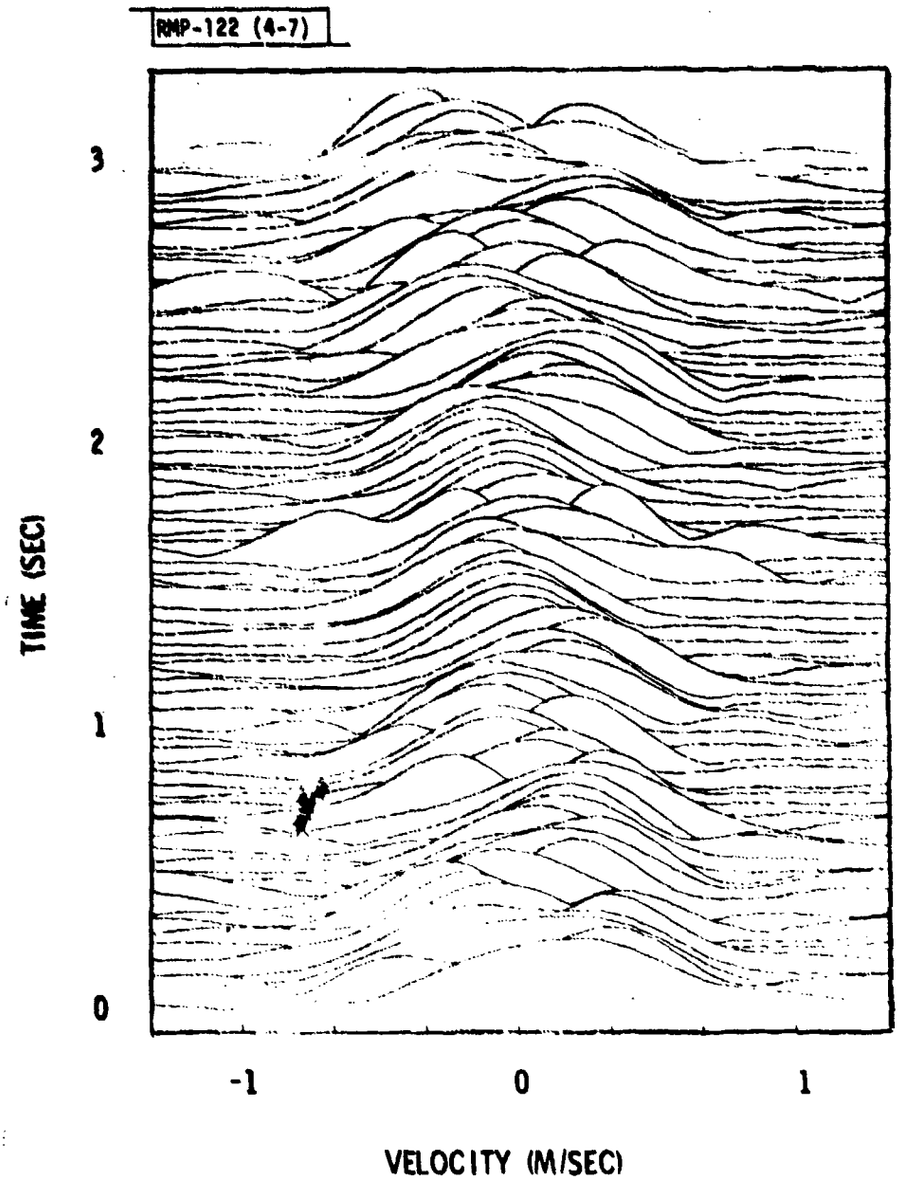


Fig.4-7. Similar to Fig.4-4, but only eight pulses are Fourier transformed at a time.

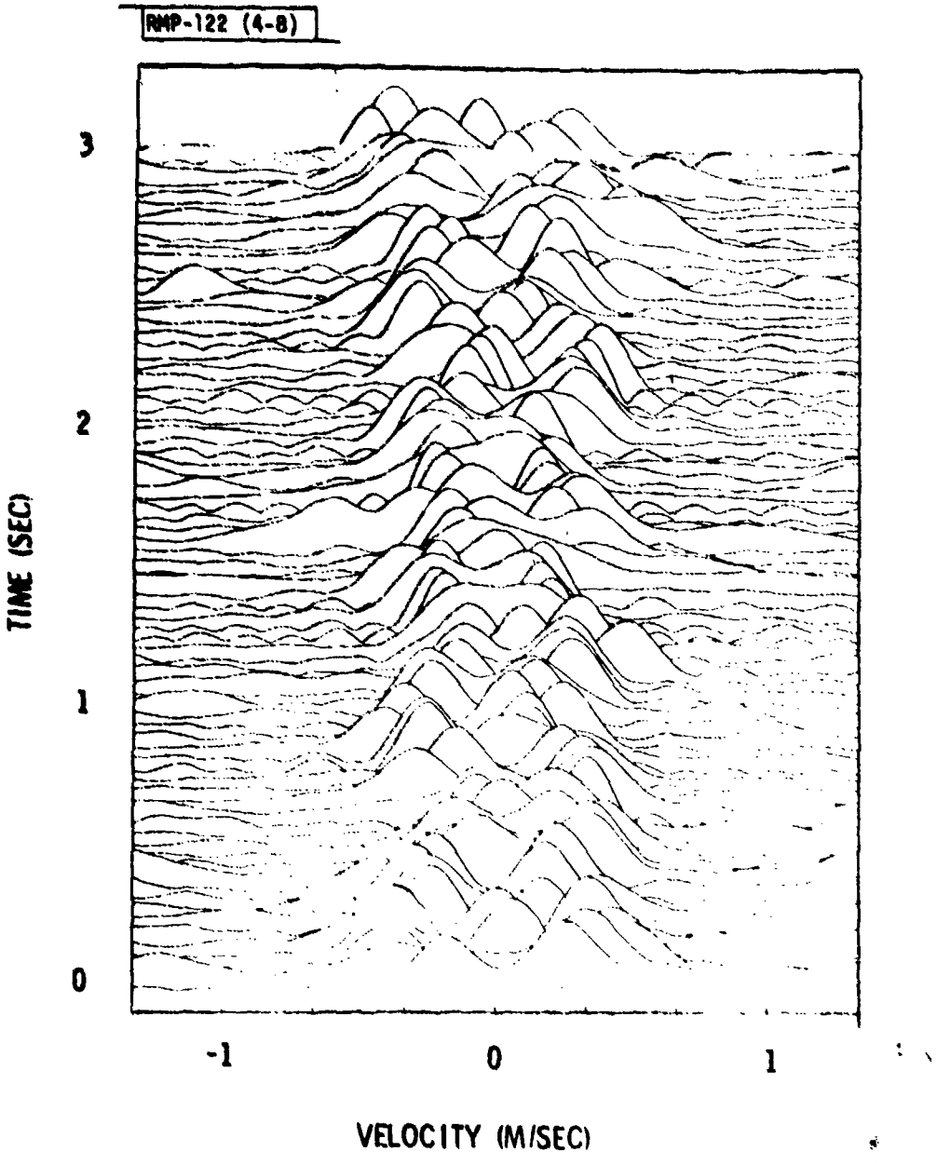


Fig.4-8. Doppler history similar to Fig.4-7, but data are linearly extended before Fourier transformation.

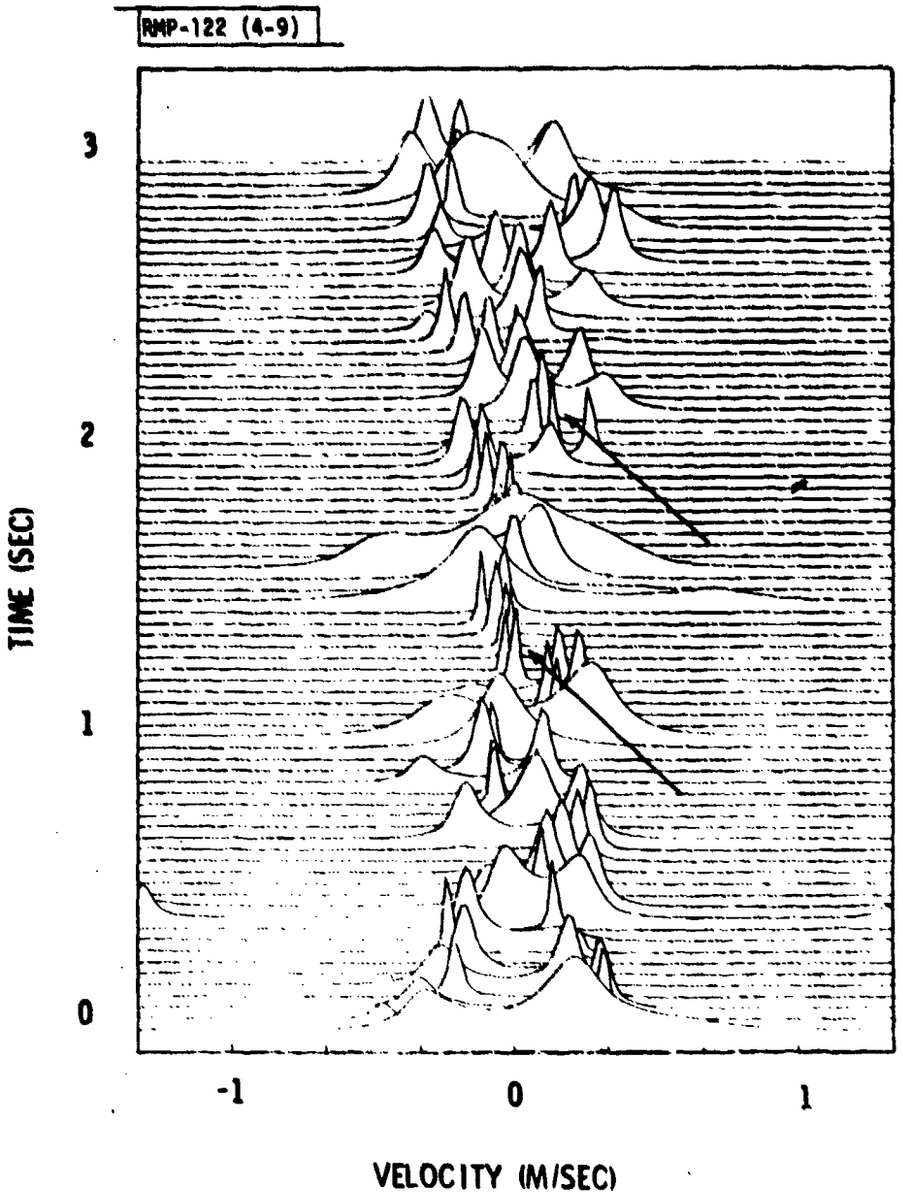


Fig.4-9. MESA Doppler history analogous to Fig.4-7, Precessional envelope becomes apparent, and tracks of individual scattering centers are perceptible (arrows).

APPENDIX I

CALCULATION OF PREDICTION-ERROR FILTER COEFFICIENTS FOR COMPLEX DATA

Van Den Bos (1971) has shown that MESA is equivalent to a least-squares all-pole model of the data being analyzed. This means that a data point is predicted by the weighted average of its neighbors. The all-pole model is also known as the autoregressive model (Box and Jenkins, 1970). Anderson (1974) used this formalism to develop algorithms for calculating the prediction filter coefficients for a real input time series. Since a complex time series is often of interest, we shall modify Anderson's work here and extend it to complex form. The same notation is used for ease in comparison. The reader should note that Anderson defines his filter coefficients as $-a_{m,n}$ instead of $+a_{m,n}$ (the latter is the convention in most MESA literature). Then, the MESA power spectrum is written

$$P(f) = \frac{P_m \Delta t}{\left| 1 - \sum_{n=1}^{m} a_{m,n} \, e^{-2\pi f n \Delta t} \right|^2} \qquad (I\text{-}1)$$

The frequency f is limited to the Nyquist range $|f| \leq (1/2\Delta t)$. P_m (the residual error power remaining after an (m+1)-point

filter is convolved with the data) and each filter coefficient $a_{m,n}$ (the nth coefficient out of the total m coefficients) are determined by the solution of the equation

$$\begin{bmatrix} \phi(0) & \phi(1) & \phi(m) \\ \phi(-1) & \phi(0) & \phi(m-1) \\ . & . & . \\ . & . & . \\ \phi(-m) & \phi(-m+1) & \phi(0) \end{bmatrix} \begin{bmatrix} 1 \\ -a_{m,1} \\ . \\ . \\ -a_{m,m} \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ . \\ . \\ 0 \end{bmatrix} \qquad \text{(I-2)}$$

where $\phi(\ell)$ is the value of the autocorrelation function at lag ℓ. For $\ell=0$, we have

$$P_o = \frac{1}{N} \sum_{t=1}^{N} x_t x_t^* \qquad \text{(I-3)}$$

as the (real) variance of the complex series $\{x_t\}$ of N points.

The solution of (I-2) involves the stepwise increase of the matrix dimension by one and the determination of (m+2) unknowns as the known autocorrelation function is being calculated. These unknowns include the m filter coefficients, the next value of the autocorrelation function, and the error power: ($a_{m,1}$, ... , $a_{m,m}$, $\phi(m)$, P_m). There are only (m+1) equations in Eq. (I-2), so an additional relation is necessary. Burg (1968) suggested that the additional relation is the minimization of the total error power (the sum of the forward error power and backward error power). For example, the forward and backward error power for a two-point

prediction-error filter $(1, a_{1,1})$ are

$$e_f^2 = \frac{1}{2(N-1)} \sum_{t=1}^{N-1} |x_{t+1} - a_{1,1}x_t|^2 \qquad \text{(I-4)}$$

$$e_b^2 = \frac{1}{2(N-1)} \sum_{t=1}^{N-1} |x_t - a_{1,1}^* x_{t+1}|^2 \qquad \text{(I-5)}$$

Minimizing the sum $e_f^2 + e_b^2$ with respect to $a_{1,1}$ gives

$$a_{1,1} = 2 \sum_{t=1}^{N-1} x_t^* x_{t+1} \Bigg/ \sum_{t=1}^{N-1} (x_t x_t^* + x_{t+1} x_{t+1}^*). \qquad \text{(I-6)}$$

In general, for an (m+1)-point filter (1 with the m coefficients),

$$e_f^2 = \frac{1}{2(N-m)} \sum_{t=1}^{N-m} \left| x_{t+m} - \sum_{k=1}^{m} a_{m,k} x_{t+m-k} \right|^2 \qquad \text{(I-7)}$$

$$e_b^2 = \frac{1}{2(N-m)} \sum_{t=1}^{N-m} \left| x_t - \sum_{k=1}^{m} a_{m,k}^* x_{t+k} \right|^2 \qquad \text{(I-8)}$$

The sum $e_f^2 + e_b^2$ is minimized with respect to $a_{m,m}$. Adding a new coefficient will require editing the old coefficients (k=1, ... , m-1) by the rule

$$a_{m,k} = a_{m-1,k} - a_{m,m} a_{m-1,m-k}^* \qquad \text{(I-9)}$$

to update the filter. Equation (I-9) is the result of the Levinson recursion relations (Wiener, 1949; Robinson, 1967) for the

solution of Toeplitz matrix equations like Eq. (I-2). The updated error power becomes

$$P_m = P_{m-1}(1 - a_{m,m}a^*_{m,m}). \tag{I-10}$$

If we set $a_{m,o} = -1$ and $a_{m,k} = 0$ for $k > m$, we can rewrite Eqs. (I-7) and (I-8) as

$$e_f^2 = \frac{1}{2(N-m)} \sum_{t=1}^{N-m} \left| \sum_{k=o}^{m} a_{m,k}x_{t+k} \right|^2 \tag{I-11}$$

$$e_b^2 = \frac{1}{2(N-m)} \sum_{t=1}^{N-m} \left| \sum_{k=o}^{m} a^*_{m,k}x_{t+m-k} \right|^2 \tag{I-12}$$

Inserting Eq. (I-9) into Eq. (I-11) and Eq. (I-12) we obtain

$$e_f^2 = \frac{1}{2(N-m)} \sum_{t=1}^{N-m} |b'_{m,t} - a_{m,m}b_{m,t}|^2 \tag{I-13}$$

$$e_b^2 = \frac{1}{2(N-m)} \sum_{t=1}^{N-m} |b_{m,t} - a_{m,m}b'_{m,t}|^2 \tag{I-14}$$

where we have defined the series

$$b_{m,t} = \sum_{k=o}^{m} a^*_{m-1,k}x_{t+k} = \sum_{k=o}^{m} a^*_{m-1,m-k}x_{t+m-k} \tag{I-15}$$

$$b'_{m,t} = \sum_{k=o}^{m} a_{m-1,k}x_{t+m-k} = \sum_{k=o}^{m} a_{m-1,m-k}x_{t+k}. \tag{I-16}$$

The alternate forms of $b_{m,t}$ and $b'_{m,t}$ are obtained by changing the index k to m-k and reversing the order of summation (commutativity). Minimizing the sum $e_f^2 + e_b^2$ with respect to $a_{m,m}$ gives

$$a_{m,m} = 2 \sum_{t=1}^{N-m} b^*_{m,t} b'_{m,t} \Bigg/ \sum_{t=1}^{N-m} (b_{m,t} b^*_{m,t} + b'_{m,t} b'^*_{m,t}). \tag{I-17}$$

It is easy to show that

$$b_{m,t} = b_{m-1,t} - a^*_{m-1,m-1} b'_{m-1,t} \tag{I-18}$$

$$b'_{m,t} = b'_{m-1,t+1} - a_{m-1,m-1} b_{m-1,t+1} \tag{I-19}$$

so that the series $b_{m,t}$ and $b'_{m,t}$ can be constructed from their previous values as m increases by one. The initial values are

$$b_{o,t} = b'_{o,t} = x_t \tag{I-20}$$

and

$$b_{1,t} = x_t$$

$$b'_{1,t} = x_{t+1} \tag{I-21}$$

Then $b_{m,t}$ and $b'_{m,t}$ can be calculated from Eqs. (I-18) and (I-19) as m is incremented from 1 to the desired filter length M.

Clearly, the solution of (I-2) is a "bootstrap" process based on a recursion from m = 1 to M. Figure I-1 shows a flow chart for the recursive procedure that calculates the complex filter coefficients. Following Anderson's Figure 1, $b_{m,t}$ is b1

and $b'_{m,t}$ is b2. The array aa(t) is temporary storage of the filter coefficients before being updated into array a(t) by Eq. (I-9), and array P(m) is the error power updated by Eq. (I-10) as m ranges from one to its maximum specified value. Element P(M) is the final error power used in the MESA spectrum.

Figure I-2 gives a FORTRAN listing for the flow chart of Fig. I-1. The arguments of the subroutine are

NPTS	number of data samples
X	complex array of data samples
NCØEFF	number of filter coefficients to be calculated
A	complex array of coefficients
PM	real array of updated error power
PØ	initial variance
AA, B1, B2	work arrays.

The storage for all arrays must be supplied by the calling program. Inputs are NPTS, X, and NCØEFF.

This subroutine can be used with the program CPSPEC given in Fig. I-3 to calculate a MESA power spectrum. The argument names are

NCØEFF	number of filter coefficients
A	complex array of coefficients

PSPEC	array of values of power spectrum
FREQ	array of frequencies at which PSPEC is calculated
NPSPEC	number of frequencies in array FREQ
PLAST	PM(NCØEFF) from above = P_m in Eq. (I-1)
DT	sample spacing of data points.

Inputs are NCØEFF, A, FREQ, NPSPEC, PLAST, and DT. Subroutine CPSPEC is written with Anderson's negative coefficient convention (Eq. (I-1)). Again, array storage must be supplied by the calling program.

Figure I-4 lists the subroutine LNPRED by which complex data may be linearly extended. All arguments are input, but array X must be large enough to accommodate the extension. Anderson's negative coefficient convention is incorporated into the code.

N1	original number of data points
N2	number of data points after extension
X	complex data array
NCØEFF	number of coefficients
A	complex array of filter coefficients.

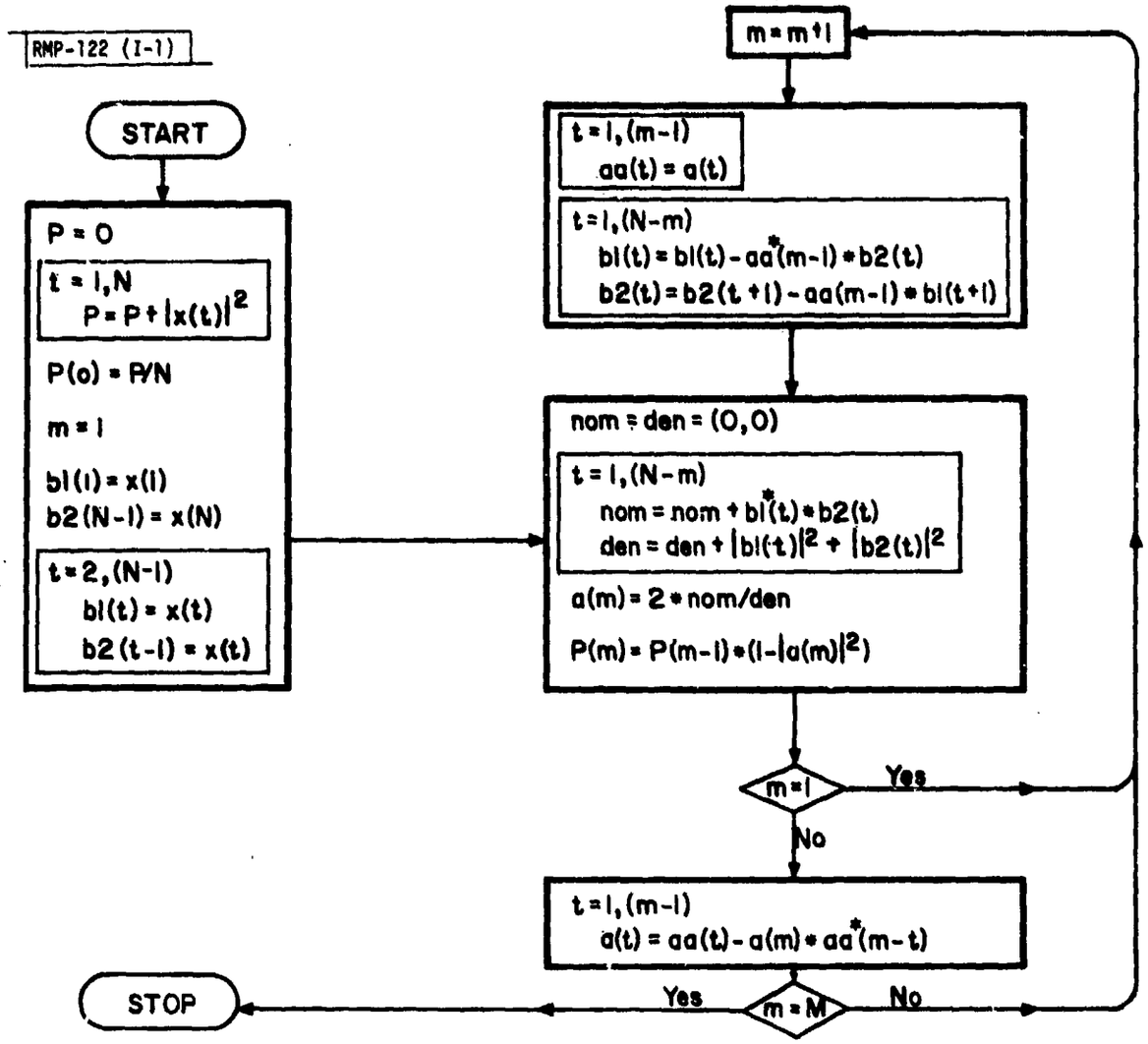


Fig.I-1. Flow chart describing calculation of complex prediction filter coefficients.

```
C
      SUBROUTINE COEFF(NPTS,X,NCOEFF,A,PM,PO,AA,B1,B2)
C
C  THIS SUBROUTINE CALCULATES THE COMPLEX FILTER COEFFICIENTS.
C  THE ALGORITHMS USED HERE ARE A MODIFICATION OF THE
C  ALGORITHMS DESCRIBED BY ANDERSON (GEOPHYSICS,VOL. 39,FEB. 1974)
C  FOR THE CASE OF A COMPLEX SERIES. REAL DATA CAN BE PROCESSED BY
C  SETTING THE IMAGINARY PART OF THE DATA STORED IN X TO ZERO.
C
C  INPUTS ARE NPTS,X,NCOEFF
C
C  NPTS   = THE NUMBER OF DATA POINTS IN THE DATA SET
C  X      = A COMPLEX ARRAY CONTAINING THE ORIGINAL DATA
C  NCOEFF = THE NUMBER OF FILTER COEFFICIENTS TO BE CALCULATED
C  A      = THE ARRAY CONTAINING THE COMPLEX FILTER COEFFICIENTS
C  PM     = REAL ARRAY CONTAINING THE UPDATED ERROR POWER
C  PO     = REAL VARIANCE OF THE DATA
C  AA, B1, B2 ARE WORK ARRAYS
C
C  STORAGE FOR THE ARRAYS MUST BE SUPPLIED BY THE CALLING
C  PROGRAM: X(NPTS),A(NCOEFF),AA(NCOEFF),B1(NPTS),B2(NPTS),PM(NCOEFF)
C  ARE THE MINIMUM STORAGE REQUIREMENTS FOR THIS SUBROUTINE.
C
C  PROGRAMMED BY S.B. BOWLING, MIT-LINCOLN LABORATORY, FEB. 1977.
C
      COMPLEX X,A,AA,B1,B2,XNOM,DEN,TWO
      DIMENSION X(1),A(1),AA(1),B1(1),B2(1),PM(1)
      TWO=CMPLX(2.0,0.0)
      PO=0.0
      DO 10 IT=1,NPTS
      DUMMY= X(IT)*CONJG(X(IT))
   10 PO= PO+ DUMMY
      PO= PO/FLOAT(NPTS)
      NM1=NPTS-1
      B1(1)=X(1)
      B2(NM1)=X(NPTS)
      DO 20 IT=2,NM1
      B1(IT)=X(IT)
      ITM1=IT-1
   20 B2(ITM1)=X(IT)
      DO 50 M=1,NCOEFF
      MM1=M-1
      NMM=NPTS-M
      IF(M .EQ. 1) GO TO 25
      DO 21 IT=1,MM1
   21 AA(IT)=A(IT)
      DO 22 IT=1,NMM
      B1(IT)= B1(IT)-CONJG(AA(MM1))*B2(IT)
   22 B2(IT)= B2(IT+1)-AA(MM1)*B1(IT+1)
   25 XNOM=CMPLX(0.0,0.0)
      DEN=CMPLX(0.0,0.0)
      DO 30 IT=1,NMM
      XNOM=XNOM + B2(IT)*CONJG(B1(IT))
   30 DEN=DEN + B1(IT)*CONJG(B1(IT)) + B2(IT)*CONJG(B2(IT))
      IF( REAL(DEN) .EQ. 0.0) GO TO 35
      A(M)= TWO*(XNOM/DEN)
      GO TO 36
   35 A(M)=CMPLX(0.0,0.0)
   36 POWER=PO
      IF(M .GT. 1) POWER=PM(M-1)
      DUMMY=A(M)*CONJG(A(M))
      PM(M)=POWER*(1.0 - DUMMY)
      IF(M .EQ. 1) GO TO 50
      DO 40 IT=1,MM1
   40 A(IT)= AA(IT)-A(M)*CONJG(AA(M-IT))
   50 CONTINUE                                          RMP-122 (1-2)
      RETURN
      END
```

Fig.I-2. FORTRAN listing for flow chart in Fig.I-1.

```
C
      SUBROUTINE CPSPEC(NCOEFF,A,PSPEC,FREQ,NPSPEC,PLAST,DT)
C
C  THIS SUBROUTINE COMPUTES THE POWER SPECTRUM PSPEC AT THE FREQUENCIES
C  STORED IN FREQ.  IT USES RECURSION RELATIONS FOR COS(N*THETA)
C  AND SIN(N*THETA) TO SAVE TIME.  IT IS CAST IN COMPLEX FORM
C  SINCE THE PREDICTION COEFFICIENTS ARE GENERALLY COMPLEX.
C
C  INPUTS ARE NCOEFF,A,FREQ,NPSPEC,PLAST,DT.
C
C  NCOEFF  = NUMBER OF FILTER COEFFICIENTS
C  A       = COMPLEX ARRAY OF FILTER COEFFICIENTS
C  FREQ    = ARRAY OF FREQUENCIES AT WHICH THE POWER SPECTRUM
C            IS TO BE CALCULATED
C  PSPEC   = ARRAY OF VALUES OF THE POWER SPECTRUM
C  NPSPEC  = NUMBER OF FREQUENCIES AT WHICH THE POWER SPECTRUM
C            IS TO BE CALCULATED = DIMENSION OF FREQ AND PSPEC
C  PLAST   = RESIDUAL ERROR POWER AFTER THE FILTER HAS OPERATED
C            ON THE DATA = ELEMENT PM(NCOEFF) FROM  SUBROUTINE 'COEFF'
C  DT      = SPACING OF THE DATA
C
C  STORAGE FOR ARRAYS IS SUPPLIED BY THE CALLING PROGRAM. MINIMUM
C  STORAGE REQUIREMENTS ARE A(NCOEFF),PSPEC(NPSPEC),FREQ(NPSPEC)
C  FOR THIS SUBROUTINE.
C
C  PROGRAMMED BY S.B. BOWLING, MIT-LINCOLN LABORATORY, FEB. 1977.
C
      COMPLEX A,ZN,DEN
      DIMENSION A(1),PSPEC(1),FREQ(1)
      FACTOR= -6.2831853*DT
      DO 100 I=1,NPSPEC
      CN=1.0
      SN=0.0
      DEN=CMPLX(1.0,0.0)
      ARG=FACTOR*FREQ(I)
      CARG=COS(ARG)
      SARG=SIN(ARG)
      DO 50 N=1,NCOEFF
      TEMP= CARG*CN - SARG*SN
      SN= CARG*SN + SARG*CN
      CN= TEMP
      ZN= CMPLX(CN,SN)
      DEN= DEN - A(N)*ZN
   50 CONTINUE
      DSQ= DEN*CONJG(DEN)
      IF( DSQ .EQ. 0.0 ) DSQ=1.0E-11
      PSPEC(I)= (1.0/DSQ)*PLAST*DT
  100 CONTINUE
      RETURN
      END
```

BEST AVAILABLE COPY

RMP-122 (I-3)

Fig.I-3. FORTRAN listing to calculate MESA power spectrum.

```
      SUBROUTINE LNPRED(N1,N2,X,NCOEFF,A)
C
C  THIS SUBROUTINE LINEARLY EXTENDS THE COMPLEX DATA  'X'
C  FROM N1 POINTS TO N2 POINTS.  THE ORIGINAL DATA , POSITIONED
C  IN THE FIRST N1 ELEMENTS OF ARRAY X, ARE SHIFTED TO THE MIDDLE OF X.
C  FORWARD AND BACKWARD PREDICTIONS ARE DONE UNTIL THE TOTAL
C  NUMBER OF DATA POINTS IS N2.
C
C  NOTE THAT THE N2 POINTS CONTAIN THE ORIGINAL N1 POINTS.
C
C  INPUTS ARE N1,N2,X,NCOEFF,A.
C
C  ARRAY X IS MODIFIED ; THE ORIGINAL DATA POINTS HAVE BEEN
C  SHIFTED TO THE MIDDLE AND PREDICTIONS ARE DONE ON BOTH SIDES.
C
C  THE FILTER COEFFICIENTS SHOULD ALREADY HAVE BEEN CALCULATED
C  BY PROGRAM 'COEFF.'
C
C  N1     = ORIGINAL NUMBER OF POINTS IN ARRAY X
C  N2     = NUMBER OF POINTS TO WHICH  ARRAY X IS TO BE EXTENDED
C  X      = COMPLEX ARRAY OF DATA SAMPLES
C  NCOEFF = NUMBER OF PREDICTION FILTER COEFFICIENTS
C  A      = COMPLEX ARRAY OF FILTER COEFFICIENTS
C
C  STORAGE FOR THE ARRAYS SHOULD BE SUPPLIED BY THE CALLING
C  PROGRAM: MINIMUM STORAGE REQUIREMENTS ARE X(N2),A(NCOEFF)
C  FOR THIS SUBROUTINE.
C
C  PROGRAMMED BY S.B. BOLLING, MIT-LINCOLN LABORATORY, FEB.1977.
C
      COMPLEX X,A
      DIMENSION X(1),A(1)
C
C  SET UP LIMITS FOR DO LOOPS
C
      L1= N2/2  - N1/2
      L2= N2/2  + N1/2
      IF( MOD(N1,2) .EQ. 1 )  L2=L2+1
C
C  SHIFT ORIGINAL DATA TO MIDDLE OF ARRAY X
C
      DO 100 I=1,N1
      J= N1 - (I-1)
      K= L2 - (I-1)
  100 X(K)=X(J)
C
C  DO FORWARD PREDICTION
C
      N3= N2-L2
      DO 200 I=1,N3
      J= L2+I
      X(J)=CMPLX(0.0,0.0)
      DO 200 K=1,NCOEFF
  200 X(J)= X(J) + A(K)*X(J-K)
C
C  DO BACKWARD PREDICTION
C
      DO 300 I=1,L1
      J= L1- (I-1)
      X(J)=CMPLX(0.0,0.0)
      DO 300 K=1,NCOEFF
  300 X(J)=X(J) + CONJG(A(K))*X(J+K)
C
      RETURN
      END
C
```

BEST AVAILABLE COPY

RMP-122 (I-4)

Fig.1-4. FORTRAN listing to linearly extend a complex data set.

APPENDIX II

A COMPARISON OF SOME OF THE CHARACTERISTICS OF MESA AND CONVENTIONAL FOURIER SPECTRAL ANALYSIS

Here we tabulate some of the salient features of the two methods of spectral analysis. For a more mathematical presentation, the reader is referred to Lacoss (1971) and to Chen and Stegun (1974). The notation in the table is

N = number of values of known autocorrelation function

P = power of a pure tone in the spectrum

Δt = spacing of input data

NPTS = number of data samples.

TABLE II-1

COMPARISON OF MESA AND FFT CHARACTERISTICS

Feature	Fourier	MESA
(1) Estimation of autocorrelation function	Zero extension of autocorrelation function; the tapering window that surpresses sidelobes modifies known values of autocorrelation function.	Known autocorrelation function unmodified; unknown values extended with linear prediction algorithm in optimal way.
(2) Spectral window effects	Convolution of window spectrum with true signal spectrum reduces resolution and allows leakage through window sidelobes.	No window effects since autocorrelation function is estimated for all lags. Minor effects occur if P-E filter does not create perfectly white output spectrum.
(3) Linearity of spectrum	Spectral estimation is linear; spectrum of sum of signals is sum of their individual spectra.	Spectral estimation is not linear; relative power amplitudes of component frequencies may not always be reliable.
(4) Peak power density of a pure tone	Proportional to P.	Proportional to $(PM)^2$.
(5) Spectral reliability	Number of degrees of freedom is 2(NPTS)/M for Bartlett window.	Difficult to define, but number of degrees of freedom is not greater than NPTS/M.
(6) Bandwidth of pure tone	Proportional to 1/M.	Proportional to $1/(PM)^2$.
(7) Accuracy of frequency estimation of pure tone	To within $\pm 1/2M\Delta t$.	Not easily defined, but can be estimated very closely.
(8) Resolution of closely-spaced frequency components	Proportional to 1/M.	Data-dependent, but resolution approximately proportional to $1/M^2$.
(9) Line detectability in spectrum	MESA spectrum always superior to Fourier spectrum for weak signals imbedded in a short data set with low processing gain. For longer data samples, the MESA spectrum improves because more accurate values of known autocorrelation function can be calculated.	

APPENDIX III

ON INTEGRATING THE MESA POWER SPECTRUM TO FIND THE POWER SPECTRAL DENSITY

Lacoss (1971) has pointed out that it is the MESA power spectral density which is the more appropriate measure of the relative power of spectral components. Since the power spectrum P(f) is calculable at any frequency $|f| \leq 1/2\Delta t$ (assuming the data are band-limited and properly sampled), in principle it is easy to perform a numerical integration over a small bandwidth δf to calculate a power spectral density. However, some of the peaks in P(f) may be so narrow and sharp that they may be difficult to detect if the initial grid spacing in frequency is coarse.

Radoski _t al._ (1975) emphasize that if the signal-to-noise ratio is high, a coarse frequency grid can be insufficient to determine the actual locations of spectral peaks in the MESA power spectrum. As the number of filter coefficients increases, the peaks tend to approach line spectra (delta functions). Radoski _et al._ (1975) have suggested a systematic search procedure (used in this report in a modified form) to calculate the power spectral density from the MESA power spectrum:

(1) The data are normalized to unit variance. Each datum x is replaced by $(x - \bar{x})/\sigma$ where $\bar{x}$ and σ are the (complex) average value and (real) standard deviation of

the original data. By Parseval's Theorem, the integral of the power spectral density (PSD) must be unity. Therefore we can feel confident that the PSD has been correctly estimated, the spectrum integrated, and all spectral lines detected, if numerically $\int PSD(f)df \approx 1$.

(2) Triplets of points in the power spectrum are examined such that, for $(f_{i+1} - f_{i-1}) = \delta f$, whenever

$$P(f_{i-1}) < P(f_i) > P(f_{i+1}) \qquad \text{(III-1)},$$

the midpoint is near a possible spectral peak. P(f) at two intermediate points is calculated and the five are examined to extract a new triplet. Simpson's rule is used to integrate P(f) as the search procedure works its way up into the peak until two calculations of the integral over the bandwidth δf are within 1% of each other. Then the procedure shifts to the next triplet of points in the power spectrum spanning δf and the integration process is repeated. If Eq. (III-1) is not satisfied, Simpson's rule integrates P(f) over δf with enough points so that two successive integrations are within 1% of each other. See Radoski et al. (1975) for details.

Normalizing the data to unit variance removes the dc or zero frequency part of the spectrum, which may not be appropriate for

some applications. If the data are not normalized, then $\int PSD(f)df$ should equal the variance σ^2 of the data (which is equal to the zero-lag value of the autocorrelation function). Depending on the form of Parseval's Theorem used, $\int PSD(f)df$ and σ^2 may differ by a factor of 2π or $(2\pi)^{1/2}$. (Parseval's Theorem states that the total average power (or mean-square value) of x(t) is equal to the integral of the power spectral density over all relevant frequencies.)

REFERENCES

Ables, J.G., "Maximum Entropy Sepctral Analysis," Astron. Astrophys. Suppl. 15, 383 (1974).

Akaike, H., "Fitting Autoregressive Models for Prediction," Ann. Inst. Statist. Math. 21, 243 (1969a).

Akaike, H., "Power Spectrum Estimation through Autoregressive Model Fitting," Ann. Inst. Statist. Math. 21, 407 (1969b).

Akaike, H., "Statistical Predictor Identification," Ann. Inst. Statist. Math. 22, 203 (1970).

Anderson, M., "On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis," Geophysics 39, 69 (1974).

Baggeroer, A.B., "Confidence Intervals for Regression (MEM) Spectral Estimates," IEEE Trans. Inf. Theory IT-22, 534 (1976) DDC AD-A037628.

Blackman, R.B. and J.W. Tukey, The Measurement of Power Spectra (Dover Publications, New York, 1959).

Box, G.E. and G.H. Jenkins, Time Series Analysis: Forecasting and Control (Holden-Day, San Francisco, 1970).

Burg, J.P., "Maximum Entropy Spectral Analysis," paper presented at 37th Annual Meeting, Society of Exploration Geophysicists, Oklahoma City, Oklahoma, 31 October 1967.

Burg, J.P., "A New Analysis Technique for Time Series Data," paper presented at Advanced Study Institute on Signal Processing, NATO, Enschede, Netherlands, 1968.

Burg, J.P., "Maximum Entropy Spectral Analysis," Ph.D Thesis, Department of Electrical Engineering, Stanford University, Xerox University Microfilms No. 75-25-499, Ann Arbor, Michigan (May 1975).

Chen, W.Y. and G.R. Stegun, "Experiments with Maximum Entropy Power Spectra of Sinusoids," J. Geophys. Res. 79, 3019 (1974).

Claerbout, J.F., Fundamentals of Geophysical Data Processing (McGraw-Hill, New York, 1976).

Edward, J.A. and M.M. Fitelson, "Notes on Maximum Entropy Processing," IEEE Trans. Inf. Theory IT-19, 232 (1973).

Jackson, P.L., "Truncations and Phase Relationships of Sinusoids," J. Geophys. Res. 72, 1400 (1967).

Jaynes, E.T., "Information Theory and Statistical Mechanics," Phys. Rev. 106, 620 (1957).

Jones, R.H., "A Reappraisal of the Periodogram in Spectral Analysis," Technometrics 7, 531 (1965).

Lacoss, R.T., "Data Adaptive Spectral Analysis Methods," Geophysics 36, 661 (1971) DDC AD-734104.

Makhoul, J., "Linear Prediction: A Tutorial Review," Proc. IEEE 63, 561 (1975).

Parzen, E., "Multiple Time Series Modeling," in Multivariate Analysis 2, edited by P.R. Krishnaiah (Academic Press, New York, 1969).

Peacock, K.L. and S. Treitel, "Predictive Deconvolution: Theory and Practice," Geophysics 34, 155 (1969).

Radoski, H.R., P.F. Fougere, and E.J. Zawalick, "A Comparison of Power Spectral Estimates and Applications of the Maximum Entropy Method," J. Geophys. Res. 80, 619 (1975).

Robinson, E.A., Statistical Communication and Detection (Hafner Publishing Co., New York, 1967).

Shannon, C.E. and W. Weaver, The Mathematical Theory of Communication (University of Illinois Press, Urbana, 1949).

Toman, K., "The Spectral Shifts of Truncated Sinusoids," J. Geophys. Res. 70, 1749 (1965).

Ulrych, T.J. and T.N. Bishop, "Maximum Entropy Spectral Analysis and Autoregressive Decomposition," Rev. Geophys. Space Phys. 13, 183 (1975).

Van Den Bos, A., "Alternative Interpretation of Maximum Entropy Spectral Analysis," IEEE Trans. Inform. Theory IT-17, 493 (1971).

Wiener, N., Extrapolation, Interpolation, and Smoothing of Stationary Time Series (Wiley, New York, 1950).

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ESD-TR-77-213	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE *(and Subtitle)*: Linear Prediction and Maximum Entropy Spectral Analysis for Radar Applications	5. TYPE OF REPORT & PERIOD COVERED: Project Report
6. PERFORMING ORG. REPORT NUMBER: RMP-122	
7. AUTHOR(s): Stephen B. Bowling	8. CONTRACT OR GRANT NUMBER(s): F19628-76-C-0002
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Lincoln Laboratory, M.I.T., P.O. Box 73, Lexington, MA 02173	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Project No. 8X363304D215
11. CONTROLLING OFFICE NAME AND ADDRESS: Ballistic Missile Defense Program Office, Department of the Army, 5001 Eisenhower Ave., Alexandria, VA 22333	12. REPORT DATE: 24 May 1977 13. NUMBER OF PAGES: 74
14. MONITORING AGENCY NAME & ADDRESS *(if different from Controlling Office)*: Electronic Systems Division, Hanscom AFB, Bedford, MA 01731	15. SECURITY CLASS. *(of this report)*: Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

radar range rate
maximum entropy
linear prediction
autoregressive decomposition
predictive deconvolution
data adaptive
cross-range sizing
Doppler image

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

For most applications in radar data processing, the Fourier transform performs satisfactorily. However, other methods of spectral analysis can offer some advantages when a data set is too short for a Fourier transform to resolve or detect important spectral features. This report describes one alternative technique, maximum entropy spectral analysis (MESA), and suggests possible radar applications including range-Doppler sizing and the coherent measurement of range rate. Practical examples demonstrate an improvement in velocity resolution and cross-range resolution. Computer codes are listed that calculate MESA power spectra for a real or complex discrete time series.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

207650